INVESTIGATION OF SITE-SPECIFIC CHANGES WITHIN THE N-TERMINAL DOMAIN OF CALMODULIN AND THE RELATIONSHIP TO STRUCTURAL AND THERMAL STABILITY

By

Talia Thresa Martin

Submitted to the graduate degree program in Pharmaceutical Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science.

_____

Chairperson Jennifer S. Laurence Ph.D.

_____

Teruna Siahaan Ph.D.

_____

Valentino J. Stella Ph.D.

Date Defended: May 13, 2011

The Dissertation Committee for Talia Thresa Martin
certifies that this is the approved version of the following dissertation:

INVESTIGATION OF SITE-SPECIFIC CHANGES WITHIN THE N-TERMINAL DOMAIN OF CALMODULIN AND THE RELATIONSHIP TO STRUCTURAL AND THERMAL STABILITY

_____

Chairperson Jennifer S. Laurence Ph.D.

Date Approved: May 13, 2011

Abstract

The goal of this dissertation was to develop a method to assess how solution conditions affect a protein's stability, and a well-characterized model protein, was used to accomplish this objective. Generally, stabilizing conditions for proteins are identified from screening a matrix of conditions using an array of biophysical methods that probe physical stability by determining melting temperatures (T_m) of secondary and tertiary structure. Although this approach has been invaluable for quickly obtaining stable formulations, these tools do not provide site-specific information, which is needed to gain mechanistic understanding of protein stability. Our methodology employs high-resolution solution NMR spectroscopy to acquire molecular information about how specific residues within the structure are impacted by solution conditions and how the conditions affect stability. The N-terminal domain of Calmodulin (N-CaM) was recombinantly expressed in ^{15}N media along with supplemental calcium to enhance *in vivo* stability. Solution NMR spectroscopy was used to evaluate the role of site-specific changes in mobility and structure on the overall stability of this model system in different solution conditions. The ^{1}H-^{15}N HSQC experiment was used to investigate effects of pH and ionic strength on individual residues within or near the two homologous Ca^{2+}-binding sites. Circular dichroism (CD) was also employed to assess the thermal stability of N-CaM and to help interpret the observed site-specific changes in ^{1}H-^{15}N HSQC spectra.

To increase protein yield and produce a sufficient amount of protein for NMR analysis, varying amounts of calcium concentrations were screened in the expression media. Based on densitometric analysis of SDS-PAGE, 1.5 mM $CaCl_2$ was determined to provide the optimal concentration for enhancing *in vivo* stability, leading to the highest yield of N-CaM expression. From 2D NMR experiments, quantitative information that reflects changes in dynamics was

derived from changes in peak shape and chemical shift position to elucidate how the solution environment affects residues within N-CaM. ^{1}H-^{15}N HSQC experiments provided quantitative determination of specific residues near known degradation sites that undergo physical perturbations. Changes in the chemical shift position and peak shape observed during ^{1}H-^{15}N HSQC experiments confirmed that increasing pH and ionic strength affects a limited subset of residues within a localized region. Residues N60, G61, and D64 in the Ca^{2+}-binding loop site I was affected, in increasing pH and high ionic strength, while residues in the corresponding position of Ca^{2+}-binding loop site II remained unaffected. CD data confirmed that, at basic pH, increased ionic strength resulted in increased overall thermal stability.

Our study used common methods of probing physical stability to evaluate the overall thermal stability of N-CaM in different solution conditions. Analysis by NMR spectroscopy provided more detailed, site-specific information to show that different solution conditions affect individual residues differently, even when the amino acid sequence and structure are highly similar. By utilizing NMR spectroscopy combined with CD analysis, we were able to show that a subset of individual residues play a role in overall stability. Employing 2D NMR experiments to provide molecular information can be utilized in evaluating solution conditions for formulation of proteins of therapeutic interest.

Since I am the first college graduate of my family, I dedicate my entire education and achievements to the second generation of my family, Akaila, Nidea, Sincere, Aarylyn, Teagan, and Kennion.

ACKNOWLEDGEMENTS

The most important part of me is “nean naneewi” my family, encompassing my mother, siblings and the long lineage of people preceding my grandfather Sato Sudow “Barney” Yokoyama and nean gagu’, my grandmothers family, The Weisers. I am proud to dedicate my hard work beginning in fall of 2001, leading to my master’s dissertation to my supportive family. They have inspired and motivated me to begin and continue on a path committed to life-long learning and an education in the science. Without my family’s sacrifices I would not have the strength to pursue a higher education and endure the difficulties of being so far from the people and the culture I love. I also dedicate this to my partner, Savannah Joe, who is also a part of my family and could not have been successful without her positive motivation during graduate school, I thank you so much for your friendship and support.

This journey has also been fulfilling because of the relationships I have built with my peers and mentors. From the beginning of my education, I have benefitted from many people, with big hearts and the willingness to drive students like myself that do not have many opportunities to obtain a higher degree before college. In particular, my graduate education could not have been possible without the funding support I received from the Madison and Lila Self Graduate Fellowship and Development Program. I was very fortunate to receive this funding and have the opportunity to have meaningful experiences with intelligent and wonderful people who I know will make a difference in society. There is a long list of mentors that have made an impact in my life, but I am appreciative to all those wonderful people and I hope that I can be as successful as they were in helping others reach their goals.

The best part of my experience in graduate school is being a part of Jennifer Laurence’s lab in pharmaceutical chemistry department. I consider Jen as one of my greatest mentors

because of her genuine caring nature and the amount of time she invested into improving my talents and strengths as a graduate student. I am also grateful to the Laurence lab members I have had the pleasure working with, learning from, and having a great time inside and outside of lab! In particular, Andi for being patient and always helpful, Natalie, Mary, and Robert for trusting me and helping me to become aware of my own strengths and contributions. From these experiences, I believe we do nothing and we go nowhere without the hard work and passion from those who truly care and understand who you are. Oose (thank you)!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page Left Intentionally Blank

CHAPTER 1

UTILIZING A MODEL PROTEIN TO INVESTIGATE MECHANISMS OF PROTEIN INSTABILITY

1.1 Introduction

The goal of the research described herein is to develop methods using solution phase nuclear magnetic resonance (NMR) spectroscopy to facilitate the investigation of mechanisms of stability, to ultimately improve protein stability. Currently used methods cannot provide site-specific structural and chemical information about intact proteins. Utilizing NMR spectroscopy to study the stability of proteins can fill this void in research by providing site-specific information for probing and understanding both chemical and physical structures of proteins. Existing knowledge of *in vivo* and *in vitro* stability, along with current technologies, can complement higher resolution tools like NMR spectroscopy to understanding the mechanisms of common protein degradation routes. Understanding these mechanisms of degradation will enable the ability to predict stability better and stabilize proteins against degradation, facilitating design of safe and effective protein-based pharmaceuticals.

Numerous biophysical and analytical tools facilitate the identification of factors that influence the *in vivo* and *in vitro* physical and chemical stability of proteins (*1*). Commonly, studies involve combinations of tools to understand degradation and the impact of conditions on overall protein stability. General approaches to investigating chemical and physical degradation involve combinations of low-resolution spectroscopic and calorimetric tools. Structural and thermodynamic information acquired from these tools contributes to analyzing overall stability, however, the relationship between the two types of degradation is still not well-understood.

Unfortunately, most tools are limited in that they cannot provide molecular information to investigate mechanisms of degradation of proteins.

The chemical and physical stability of protein pharmaceuticals is an extremely important, but complex aspect of the production and formulation of therapeutic proteins. Table 1.1 displays some of the degradation routes and factors that are associated with protein stability. Generally, protein degradation is classified as either chemical or physical degradation. Chemical degradation occurs when covalent bonds are made or broken, resulting in the modification of the amino acid sequence, peptide bonds, or disruption of native covalent interactions (*2*). Degradation products from oxidation and hydrolysis related reactions (deamidation and aspartate racemization) can result in contaminants, reduced bioactivity, and oligomerization/aggregation (*3*). Physical degradation often appears as denaturation, precipitation, and/or aggregation, also resulting in contaminants and degradation products. Minimizing degradation products encountered during production and formulation of therapeutic proteins is a constant concern because it can result in lost product, reduced efficacy, and/or an immunogenic response through which overall safety is jeopardized (*2, 3*).

Current approaches to analyzing chemical and physical degradation require combinations of chemical, biophysical, and analytical tools. Spectroscopic and calorimetric instruments are used to observe physical changes and obtain melting temperatures (T_m) that reflect the state of the protein upon unfolding (*4*). Chemical modifications can be identified from changes in amino acid sequence and charge using mass spectrometry, chromatography, and sodium dodecyl sulfate polyacrylamide gels (*5*). Physical alterations to the secondary and tertiary structure are monitored by absorbance and fluorescence of intrinsic and extrinsic chromophores and fluorophores, respectively. Although these tools are helpful in providing a rapid analysis for

studying chemical and physical instability, it is cumbersome to obtain molecular information from these methods. Utilizing a higher resolution tool such as NMR spectroscopy to study protein stability can obtain molecular information about both the chemical and physical states of a protein without destruction or manipulation of the sample (*6*).

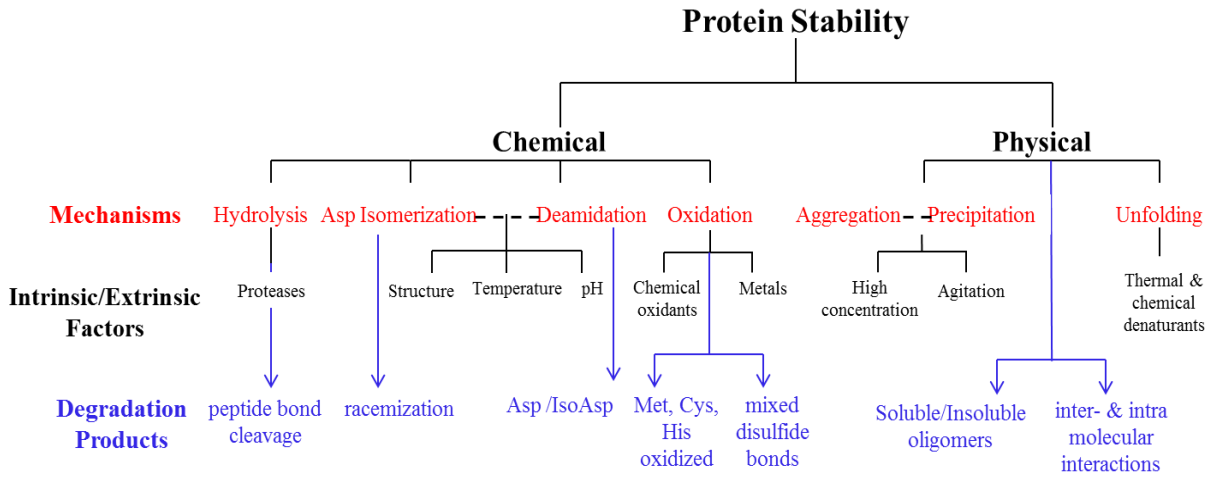

Table 1.1 **Relationships between different types of chemical and physical degradations**. Types of chemical and physical degradation leading to protein instability (red) and examples of factors influencing them (black). Degradation products resulting from each type of degradation (blue). Dashed lines (- - -) connects degradation process that have similar mechanisms and/or influences.

1.2 Techniques for Analyzing Protein Stability

Degradation of proteins is generally classified as either chemical or physical degradation. The impact of either of these two types of degradations can influence a therapeutic protein's bioactivity and/or safety. Therefore, it is important to detect products of degradation. It is a common practice and sometimes required to combine analytical and biophysical techniques to assess protein structure and chemical modifications in different solution conditions. Table 1.2 outlines commonly used tools, for analyzing protein structure and different mechanisms of degradation.

Generally, studies investigating physical and chemical stability involve using chromatographic, electrophoretic, and spectroscopic tools. To obtain a protein's chemical or structural information in different solution conditions, the method of choice is normally based on the unique properties or inherent stability of the particular protein under study. Utilization of these types of techniques along with developing new methods for protein analysis is required to understand the mechanisms of degradation and how chemical and physical degradation impact one another.

1.2.1 Chemical Stability Analysis

Chemical stability of proteins is characterized by modifications in the primary sequence or bonding between individual atoms or residues. Types of chemical degradation that can lead to a change in molecular weight and/or charge may be detected by electrophoretic methods, such as isoelectric focusing (IEF) or sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Disulfide bond formation and reduction are chemical modifications that results in changes in charge and mass. These changes can affect shape and/or hydrophobicity which can affect the mobility of the protein through a matrix, specifically for non-reducing SDS-PAGE (*5*) and Native PAGE gels. One limitation to electrophoresis is the diminished ability to detect subtle changes in chemical structure with increasing size of the protein. To compensate for poor resolution of electrophoretic techniques, mass spectrometry (MS) is used to identify subtle changes in mass of peptides and large proteins (*5, 7*). Size also diminishes resolution in MS and makes it difficult to analyze an intact, higher molecular weight protein. A well-resolved spectrum from a large molecular weight protein is therefore, difficult to achieve (*7*). Because of the size limitation, it is often necessary to employ an enzymatic or chemical digestion and/or

Edman degradation method to analyze smaller fragments. HPLC is typically coupled with MS to improve separation and to observe subtle chemical alterations. This is a common method employed by numerous studies to detect deamidation in proteins such as hGH and CaM (*5, 8-10*).

Digestion facilitates the analysis of the degradation products of deamidation (aspartate an iso-aspartate) by providing smaller fragments of the protein for analysis. In conditions that promote deamidation, the formation of iso-aspartate (Iso-Asp) can only be distinguished from Asp by using a second analysis tool. For example, protein carboxylmethyltranferase (PIMT) is used to catalyze the methylation of Iso-Asp in order to identify it by MS-HPLC (*8*). Utilization of digestions and mass spectrometry involves the destruction of sample and consequently the sample cannot be analyzed by other methods thereafter. Additionally, this method often involves exposing the sample for extensive amounts of time to conditions such as high pH and temperature, which could lead to structural alterations along with chemical modifications. This method is laborious and is adequate for detecting Iso-Asp, but it cannot ascertain information on structural state or alterations that may have occurred as a result of the chemical modification.

Chemical modifications occurring to specific residues can be obtained by utilizing hydrogen/deuterium (H/D) exchange in combination with liquid chromatography and electrospray ionization mass spectrometry (ESI-LC/MS) for analysis (*11*). This methodology involves deuterium labeled protein and a peptic digestion performed in acidic conditions. Chemical information can be obtained and located within the primary sequence. The conditions needed for analysis can be problematic because low pH conditions can also lead to deamidation via a different mechanism and physical degradation. Again, this methodology does not provide information on the impact of chemical degradation on structural changes that may occur to the

secondary and tertiary structure and additional tools must be used for assessing physical structure.

1.2.2 Physical Stability Analysis

Investigation of the stability of secondary and tertiary structure involves analysis of protein structure in physically stressful conditions that induce denaturation, aggregation, and/or precipitation. Monitoring structural alterations as a function of temperature provides information about the thermal stability of the global structure of the protein. Chemical denaturants, such as urea and guanidine hydrochloride are regularly used to facilitate exploration of physical stability of proteins undergoing temperature titrations. Spectroscopic and calorimetric tools are commonly used to investigate structural perturbations and unfolding of a protein in denaturing conditions. Specifically, alterations in secondary structure are monitored by CD, Fourier transform infrared (FTIR) and UV-Vis absorption spectroscopy. Composition of secondary structural elements and the transition temperature (T_m) can be acquired by far-UV CD (190-250 nm) and FTIR to evaluate thermal stability of secondary structure (*12*). Perturbations in the tertiary structure are observed by monitoring the absorbance and intrinsic fluorescence of aromatic residues. Extrinsic fluorescence utilizing fluorophores, such as 8-anilino-1-naphthalenesulfonate (ANS), can provide information on whether aromatic residues are in an aqueous or hydrophobic environment. This is based on the fluorescence of ANS which fluoresces when bound to aromatic residues. Analyzing the change in the microenvironment of the aromatic residues in denaturing conditions provides information about the stability of global structure and dynamics (*13*).

Physical degradation may be analyzed by electrophoretic, spectroscopic, and

chromatographic tools. SDS-PAGE and capillary electrophoresis are used to identify oligomers and higher molecular weight contaminants and to confirm the presence of the protein product. Light scattering is capable of monitoring the onset of aggregate formation by measuring increasing size of particles in solution. Analytical ultracentrifugation (AUC) and size-exclusion chromatography (SEC) can be employed to determine molecular size and separation of oligomers, dimers, and monomers (*14*). Ion-exchange chromatography is used to separate proteins by charge, which can be of interest when there are changes that alter the isoelectric point of a protein. Combining these different types of analytical tools for monitoring aggregate formation provides a qualitative assessment for understanding the numerous mechanisms of protein aggregation. Mechanisms of protein aggregation are very complex and preventing aggregate formation is recognized as an important consideration in protein production and formulation (*3*).

The advantage of using these types of biophysical and analytical tools to probe structural and physical alterations, is their ability to provide a rapid analysis of structural stability while requiring small amounts of protein. Although these tools provide global information about the overall physical stability of a protein, they are unable to give information about local interactions. NMR spectroscopy is a valuable tool for probing protein stability because of the ability to assess global and local structure of both peptides and proteins. Multi-dimensional NMR experiments take advantage of repeating elements in peptide bonds to selectively detect nuclei and obtain data pertaining to specific residues. One experiment, in particular, selectively detects the amide bond where every proton is attached to a nitrogen in the protein backbone and side-chains. Experiments involving the detection of amide groups benefit from uniform ^{15}N-labeling of the protein due to low natural abundance of ^{15}N and moderate sensitivity of NMR. Moreover,

highly concentrated sample is desirable (>1 mM). Stability of intact protein and variants can be studied by monitoring amide groups that correspond to individual residues to assess overall structure. Based on the spectra acquired from these experiments in different solution conditions, conformational and chemical changes reveal information about the global and local stability (*6*). Molecular information generated from NMR experiments in conditions that promote chemical degradation can contribute to understanding the role of how physical perturbation of individual residues may affect chemical degradation.

ENZYMATIC	Primary Uses	Chemical	Physical
Proteases	peptide digestion	✓	
Edman Degradation	sequencing	✓	
ELECTROPHORESIS			
SDS-PAGE	molecular weight analysis	✓	
Isoelectric	charge separation	✓	
Capillary	charge separation	✓	
Native PAGE	charge and hydrodynamic size	✓	✓
CHROMATOGRAPHY			
Size-Exclusion (SEC)	separation of oligomers		✓
HPLC	separation of oligomers		✓
Ion-Exchange (IEC)	separation by charge	✓	
SPECTROSCOPY			
Circular Dichroism	secondary structure		✓
Fourier Transform Infrared	secondary structure		✓
Fluorescence	tertiary structure		✓
UV absorbance	secondary & tertiary structure		✓
Nuclear Magnetic Resonance	chemical & structural	✓	✓
OTHER			
Mass Spectrometry	molecular weight, charge	✓	✓

Table 1.2 **Tools to investigate protein stability**. This table presents different categories of tools used to study chemical and physical degradation. A check indicates whether chemical and/or physical changes may be detected.

1.3 Protein Deamidation: Human Growth Hormone and Calmodulin

There are many studies about the chemical degradation via deamidation and isomerization occurring to therapeutically relevant proteins, including human growth hormone (hGH), and the model protein calmodulin (CaM) (*2, 3, 8, 15*). The most common mechanisms of *in vivo* and *in vitro* spontaneous degradation of hGH and CaM are asparagine deamidation and aspartate racemization (Figure 1.1). The degradation pathways discussed are not unique to these proteins, but they are common to many proteins and, therefore, must be considered during the development of all protein formulations. Deamidation of asparagine (Asn) and glutamine (Gln) side-chains is one of the most major concerns of protein stability during production and formulation of a protein therapeutic. Degradation products that arise from deamidation and isomerization can compromise the safety and/or efficacy of the protein product (*3*). Knowledge of the possible factors that contribute to chemical and/or physical degradation during protein expression, isolation, and purification can facilitate optimization of the yield of the final protein product and also formulation strategies.

1.3.1 General Mechanism of Protein Deamidation

Protein deamidation in solution is one of the most prevalent non-enzymatic chemical degradation pathways, which occurs spontaneously as a function of sequence, structure, and solution conditions, primarily pH (*3, 8, 10, 16*). These same conditions also affect aspartate residues, promoting isomerization and resulting in racemization of products. Figure 1.3 shows the general mechanism of deamidation of Asn under basic conditions. Deamidation in alkaline pH is initiated by the deprotonation of the amide backbone nitrogen. A nucleophilic attack by the deprotonated backbone nitrogen on the side-chain at the carbonyl carbon forms a 5-membered

cyclic succinimide (Asu) intermediate. The subsequent hydrolysis of the intermediate results in the formation of the predominate product iso-aspartate (Iso-Asp) and some aspartate (Asp). In peptides, typically 75% Iso-Asp and 25% Asp is formed.

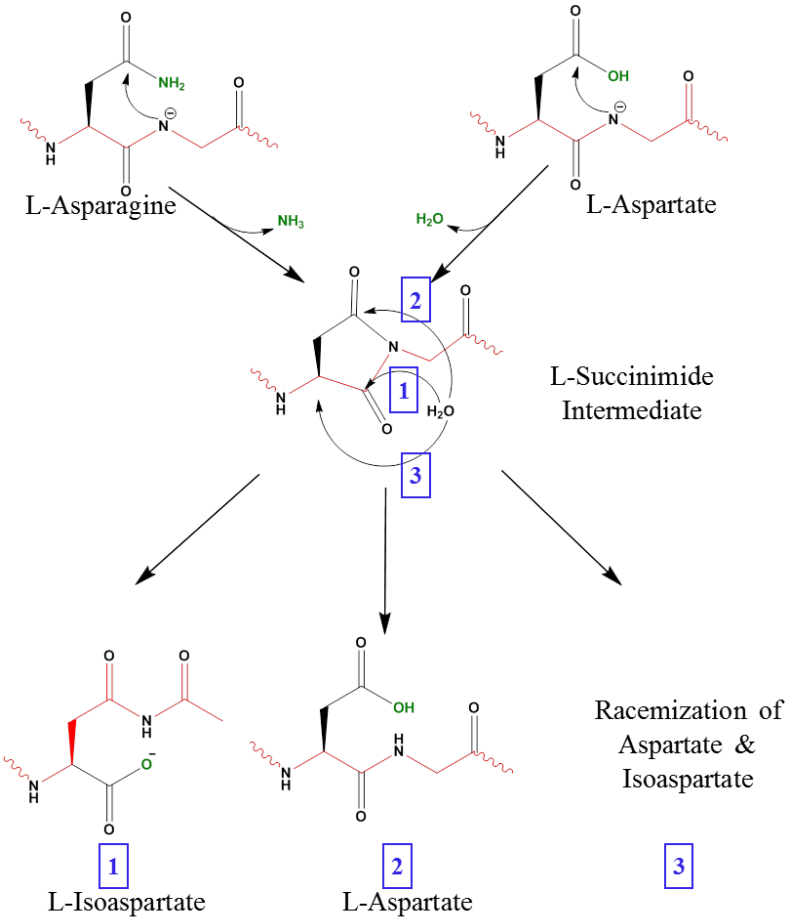

Figure 1.1 **Mechanism of deamidation and isomerization**. Changes in backbone and chemical structure is highlighted in red and green, respectively. **1.** L-Isoaspartate is the predominate degradation product formed. **2.** L-Aspartate degradation product. **3.** Mixture of racemized degradation products, L- and D- isomers.

A similar deamidation mechanism occurs with Gln, but Asn residues are known to undergo spontaneous deamidation at a substantially faster rate than Gln residues. It has been suggested, based on fundamental chemical principles, that deamidation of Asn is more rapid because 5-

membered rings formed by the Asu intermediate are relatively more stable than 6-membered rings in a protein (*2*). Another reason may be due to 5-membered rings forming intermolecular hydrogen bonds more readily than 6 membered rings (*8*).

The chemical modifications due to deamidation are a major concern for protein stability. Because deamidation can lead to protein variants with Iso-asp or Asp residues, the structural and chemical consequences may lead to an immunogenic response, inactivity, or further degradation (*16*). Moreover, the outcome is not predictable for individual proteins and the same chemical change may lead to different outcomes among different proteins. Because Asn deamidation is a predominant and rapid mechanism of degradation, the existence of Asn in proteins is highly relevant to studies in protein degradation.

The majority of the information available about the mechanism and factors contributing to deamidation derive from studies promoting deamidation on peptides under various solution conditions (*10*). These experiments focused on the susceptibility and the rate of deamidation of Asn and Gln side-chains and how they are influenced by pH, temperature, buffer, primary sequence, and ionic strength (*8, 10, 16, 17*). The pH of the solution is of primary concern because it has the largest impact on both the rate and mechanism of deamidation. The rate-limiting step at basic pH is the formation of the succinimide intermediate, which occurs increasingly rapidly above pH 6 as the backbone nitrogen become more deprotonated. In acidic conditions below pH 4, direct hydrolysis of the amine portion of the amide side-chain leads to deamidation products and racemization products, D-IsoAsp and D-Asp, although at much slower rates than occur under alkaline conditions. The residue adjacent to the Asn residue also impacts the rate and susceptibility. Less sterically hindered residues, such as glycine (Gly) and serine (Ser) at the N+1 position have been identified in labile sequences. Many charged residues at

N+1 may facilitate the mechanism by general acid and base catalysis, but to a lesser degree than Gly and Ser. General acid and base catalysis has also been observed with additives, such as phosphate, carbonate, and borate. An increase in the rate of deamidation in the presence of these buffers was observed with some peptides (*8*). Temperature has also been demonstrated to affect the rate, by increasing the flexibility of the backbone, thereby facilitating succinimide formation. Finally, the stabilization and destabilization of peptides in the presence of increased ionic strength can also affect the rate of deamidation indirectly Ionic strength alters the dielectric constant of the solvent to increase hydrophobic packing of the core of the protein which can either stabilize or destabilize tertiary and overall structure, depending on the protein (*18*).

1.3.2 Deamidation of hGH and CaM

Table 1.3 displays the sequences within hGH and CaM that are prone to deamidation in neutral to high pH. The positions in hGH identified to be most susceptible to deamidation are N149 and N152 (*8*). Deamidation at 149 and 152 does not affect hGH binding to its receptors, but conformational changes and aggregate formation are observed. There are two sites where deamidation occurs in CaM and both have an impact on Ca^{2+} binding and stability. Studies on the *in vitro* aging of CaM have identified two primary deamidation sites, one in each of the two Ca^{2+} binding sites: N97 and N60 (*9*).

CaM is an acceptable model for assessing how deamidation may impact structural degradation because of the numerous stability studies that characterize apo and Ca^{2+} bound CaM (holo CaM) (*8, 9, 19-23*). Masino *et al.* employed CD spectroscopy to show that whole CaM and isolated domains have decreased thermal stability when the Ca^{2+} ligand is absent (*24*). Additionally, several studies show that the C-domain has a higher affinity for Ca^{2+} and that it is

inherently more thermally stable than N-domain in intact CaM (*21, 24*).

Calmodulin

Binding Site	Sequence
I	$K_{21}\mathbf{D_{22}}G_{23}\mathbf{D_{24}}G_{25}T_{26}I_{27}$
II	$V_{55}D_{56}A_{57}\mathbf{D_{58}}G_{59}\mathbf{N_{60}}G_{61}$
III	$\mathbf{D_{93}}K_{94}D_{95}G_{96}\mathbf{N_{97}}G_{98}Y_{99}$
IV	$N_{129}I_{130}\mathbf{D_{131}}G_{132}\mathbf{D_{133}}G_{134}$

Human Growth Hormone

Structure	Sequence
Loop II	$F_{96}V_{97}A_{98}\mathbf{N_{99}}S_{100}L_{101}$
Loop III	$L_{128}E_{129}\mathbf{D_{130}}G_{131}S_{132}P_{133}$
Loop III	$F_{146}D_{147}A_{148}\mathbf{N_{149}}S_{150}H_{151}$
Helix III	$H_{151}\mathbf{N_{152}}D_{153}D_{154}\mathbf{A_{155}}L_{156}$

Table 1.3 Deamidation (red) and isomerization (blue) sites in the binding sites of calmodulin (left) and human growth hormone (right)

1.3.3 Physical Stability of hGH and CaM

In this work, we explore the aspects of *in vivo* and *in vitro* stability of the N-terminal domain of the model protein, calmodulin (N-CaM). Although, this dissertation focuses on N-CaM, understanding the stability of human growth hormone (hGH) has also helped to further our understanding of expressing and purifying proteins (Appendix 1). Many previous reviews and studies have contributed to understanding hGH's (*15, 25-27*) and CaM's (*19-22*) physiological roles as well as their chemical and physical stability *in vivo* and *in vitro*. Existing knowledge of these proteins facilitates recognition of conditions that induce or inhibit common routes of degradation *in vivo* and how it relates to *in vitro* stability, and we have built on this information to probe mechanism. The majority of our work focuses on the N-terminal domain of CaM. Abundant protocols for expressing and purifying CaM and a thorough characterization of CaM's thermal and chemical stability were available to build upon (*19-22*). Many studies confirm the stable and bioactive form of CaM is bound to calcium (*21*) and show that without the addition of

calcium, CaM is susceptible to *in vivo* and *in vitro* degradation. The rate of deamidation is accelerated by increasing pH or temperature, which promote deamidation of asparagines in the calcium binding sites (*23*).

Sample preparation for the analysis of protein stability begins with expression and purification of the model protein. Expressing protein from bacterial cells requires separation from cell debris and contaminants to make use of the protein. Analytical tools required to purify proteins may be used with a range of different conditions (pH, temperature, solvent), which may allow for degradation during protein isolation. Here, we chose conditions based on physiochemical properties of the protein that promote stabilization. Chromatography, SDS-PAGE, and mass spectrometry can be used to assess degradation products that emerge during purification. The *in vitro* stability can then be explored after purification by utilizing biophysical techniques, such as circular dichroism (CD) spectroscopy, to evaluate the structural state of the protein. NMR spectroscopy detects more subtle perturbations, and we utilize this capability to investigate how solution conditions influence stability at specific residues in the intact protein (*28*). The existing investigative tools, such as chromatography, electrophoresis, and CD are needed to provide a quick analysis of the material produced by the expression and purification methods. NMR experiments complete this analysis by providing assessment of structure and chemical stability relevant to the overall integrity of the protein product.

1.4 Outline of Dissertation

This dissertation demonstrates how understanding the *in vivo* and *in vitro* stability and using the current knowledge of known stabilizers can be used to optimize the expression and purification of therapeutic proteins. Chapter 2 describes the process of determining how varying

the concentration of Ca^{2+} during *in vivo* expression of N-CaM affects production yield. We analyzed the expression and final yield of N-CaM to identify the optimal concentration of Ca^{2+} that provides maximal *in vivo* stability. This chapter describes steps that enabled us to obtain a highly concentrated sample of N-CaM on which ^{1}H-^{15}N HSQC NMR experiments could be performed. NMR spectroscopy has the ability to provide more detailed information at the molecular level to further understand the mechanism of this stabilization without using destructive or numerous laborious techniques. Chapter 3 is an analysis of the NMR-based site-specific stability of N-CaM in different solution conditions. This study provides a rapid approach to obtaining qualitative and quantitative information from NMR spectra to probe protein stability. Appendix 1 contains data on a range of solution conditions used to express and purify recombinant human growth hormone from *E. coli*. In addition, an explanation of the challenges during expression and purification of growth hormone are described in Appendix 1.

1.5 References

1. Bartlett, A. I., and Radford, S. E. (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms, *Nat Struct Mol Biol16*, 582-588.
2. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W., and Katayama, D. S. (2010) Stability of protein pharmaceuticals: an update, *Pharm Res27*, 544-575.
3. Wang, W. (1999) Instability, stabilization, and formulation of liquid protein pharmaceuticals, *Int J Pharm185*, 129-188.
4. Reubsaet, J. L., Beijnen, J. H., Bult, A., van Maanen, R. J., Marchal, J. A., and Underberg, W. J. (1998) Analytical techniques used to study the degradation of proteins and peptides: physical instability, *J Pharm Biomed Anal17*, 979-984.
5. Reubsaet, J. L., Beijnen, J. H., Bult, A., van Maanen, R. J., Marchal, J. A., and Underberg, W. J. (1998) Analytical techniques used to study the degradation of proteins and peptides: chemical instability, *J Pharm Biomed Anal17*, 955-978.
6. Skinner, A. L., and Laurence, J. S. (2010) Probing Residue-Specific Interactions in the Stabilization of Proteins Using High-Resolution NMR: A Study of Disulfide Bond Compensation, *Journal of Pharmaceutical Sciences99*, 2643-2654.
7. Siuzdak, G. (2006) *The Expanding Role of Mass Spectrometry in Biotechnology* 2ed., MCC Press, San Diego.
8. Aswad, D. W., (Ed.) (1995) *Deamidation and isoaspartate formation in peptides and proteins*, CRC Press.
9. Potter, S. M., Henzel, W. J., and Aswad, D. W. (1993) In vitro aging of calmodulin generates isoaspartate at multiple Asn-Gly and Asp-Gly sites in calcium-binding domains II, III, and IV, *Protein Sci2*, 1648-1663.
10. Robinson, N. E., and Robinson, A. B. (2004) Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins., Althouse Press, Cave Junction.
11. Li, Y., Williams, T. D., Schowen, R. L., and Topp, E. M. (2007) Trehalose and Calcium Exert Site-Specific Effects on Calmoduin Conformation in Amorphous Solids *Biotechnology and Bioengineering 97*, 4.
12. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism, *Bba-Proteins Proteom1751*, 119-139.
13. Esfandiary, R., Hunjan, J. S., Lushington, G. H., Joshi, S. B., and Middaugh, C. R. (2009) Temperature dependent 2nd derivative absorbance spectroscopy of aromatic amino acids as a probe of protein dynamics, *Protein Sci18*, 2603-2614.
14. Wang, W., and Roberts, C. J., (Eds.) (2010) *Aggregation of Therapeutic Proteins*, WILEY.
15. Pearlman, R., and Bewley, T. A., (Eds.) (1993) *Stability and Characterization of Human Growth Hormone*, Plenum Press, New York.
16. Wakankar, A. A., and Borchardt, R. T. (2006) Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization, *J Pharm Sci95*, 2321-2336.
17. Stratton, L. P., Kelly, R. M., Rowe, J., Shively, J. E., Smith, D. D., Carpenter, J. F., and Manning, M. C. (2001) Controlling deamidation rates in a model peptide: effects of temperature, peptide concentration, and additives, *J Pharm Sci90*, 2141-2148.

18. Baldwin, R. L. (1996) How Hofmeister Ion Interactions affect Protein Stability *Biophysical Journal71*.
19. Biekofsky, R. R., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1998) Ca2+ coordination to backbone carbonyl oxygen atoms in calmodulin and other EF-hand proteins: 15N chemical shifts as probes for monitoring individual-site Ca2+ coordination, *Biochemistry37*, 7617-7629.
20. Jurado, L. A., Chockalingam, P. S., and Jarrett, H. W. (1999) Apocalmodulin, *Physiol Rev79*, 661-682.
21. Klee, C. B., Crouch, T. H., and Richman, P. G. (1980) Calmodulin, *Annu Rev Biochem49*, 489-515.
22. Mal, T. K., and Ikura, M. (2006) NMR Investigation of Calmodulin, *Modern Magnetic Resonance*.
23. Ota, I. M., and Clarke, S. (1989) Calcium affects the spontaneous degradation of aspartyl/asparaginyl residues in calmodulin, *Biochemistry28*, 4020-4027.
24. Masino, L., Martin, S. R., and Bayley, P. M. (2000) Ligand binding and thermodynamic stability of a multidomain protein, calmodulin, *Protein Sci9*, 1519-1529.
25. Kasimova, M. R., Kristensen, S. M., Howe, P. W., Christensen, T., Matthiesen, F., Petersen, J., Sorensen, H. H., and Led, J. J. (2002) NMR studies of the backbone flexibility and structure of human growth hormone: a comparison of high and low pH conformations, *J Mol Biol318*, 679-695.
26. Ricci, M. S., and Brems, D. N. (2004) Common structural stability properties of 4-helical bundle cytokines: possible physiological and pharmaceutical consequences, *Curr Pharm Des10*, 3901-3911.
27. Wells, J. A., and de Vos, A. M. (1993) Structure and Function of Human Growth Hormone: Implications for the Hematopoietins, *Annual Review of Biophysics and Biomolecular Structure22*, 329-351.
28. Skinner, A. L., and Laurence, J. S. (2008) High-field solution NMR spectroscopy as a tool for assessing protein interactions with small molecule ligands, *J Pharm Sci97*, 4670-4695.

Page Left Intentionally Blank

CHAPTER 2

OPTIMIZATION OF N-CAM EXPRESSION AND PURIFICATION UTILIZING A LIGAND

2.1 Introduction

The model system under analysis involves the eukaryotic Ca^{2+}-binding protein, Calmodulin (CaM). Numerous studies and reviews have contributed to the characterization of the structure and the important biological roles of CaM in various organisms (*19-22, 29, 30*). These roles are largely Ca^{2+}-dependent and facilitate the regulation of more than 30 proteins and enzymes involved in different cellular processes. These studies use a common approach to expressing CaM in bacterial cells and utilizing chromatography to purify and obtain pure, intact CaM. For studies involving analyzing CaM with multiple spectroscopic methods, a large amount of protein is required for detecting and monitoring CaM and obtaining reproducible results. Efficient use of time and resources require improving methods for expressing and purifying CaM used in these studies. One of our approaches to increasing the yield of CaM focuses on protecting CaM from physical and chemical degradation in the expression media and cellular environment. The addition of the known binding partner, Ca^{2+}, whose primary role in the media is to act as a ligand protectant when bound to CaM (holo CaM) and protect it from degradative pathways upon expression. This chapter focuses on using the Ca^{2+}, as a ligand protectant in the expression media to enhance the stability of the N-terminal domain of CaM (N-CaM), thereby improving the overall yield of pure, intact protein.

2.1.1 Optimizing the Production of N-Cam Utilizing Ligand Binding

The purpose of this analysis is to utilize existing knowledge of the stability of CaM to improve the expression and final yield of N-CaM. Because of CaM's diverse functions and biological significance, the chemical and physical stability of CaM and the isolated domains are well known. Additionally, the binding of calcium ions in the binding loops of the pair of EF hands in the N- and C-domains have been exhaustively studied by spectroscopic techniques (Figure 2.1) (*19, 22, 24, 30, 31*). The isolated domains display different thermal stabilities when they are separate or conjoined in whole CaM, and their stabilities depend on the solution conditions used to study the stability of CaM (*24, 32*). Altogether, these studies confirm the Ca^{2+}-bound state of CaM and each isolated domain is much more thermally stable and less susceptible to chemical and physical degradation compared to the apo state (Figure 2.1).

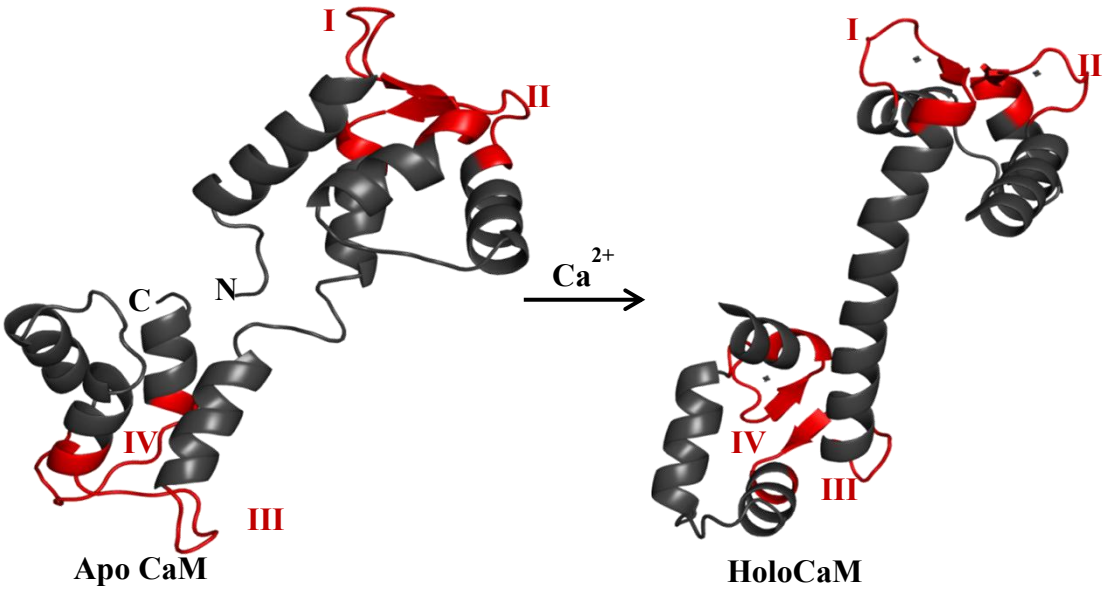

Figure 2.1 **Structures of apo and holo CaM.**
Three-dimensional structures of calmodulin (148 residues, 16.5 kD), not Ca^{2+}-bound (PDB: 3cln) and Ca^{2+}-bound (PDB: 1cfd) form of CaM. Structures are labeled with four Ca^{2+}-binding sites (I-IV) in the N and C domains. Holo form is bound to four Ca^{2+} ions.

We hypothesize that the addition of calcium during the different stages of expression and purification will improve the overall yield and integrity of N-CaM. In this case, ligand binding is performing as a protectant against degradation in chemically and physically stressful conditions, without affecting the structural integrity of the protein. Because various studies confirm Ca^{2+}-bound CaM is more thermally stable, we designate the binding of Ca^{2+} ions as a ligand protectant because of its ability to protect N-CaM from chemically degrading solution conditions and bacterial proteases during the purification and expression processes. Some studies suggest the exogenous addition of prosthetic groups and/or cofactors to optimize the production of proteins in *E. coli* (*33*). This improvement of overall yield has been observed with heme-containing proteins, in which the addition of a heme precursor was used to obtain active and a higher yield of soluble protein. Although, calcium is a cell signaling molecule, like cofactors it influences activity and has the unique advantage that it enhances CaM stability. For the purpose of optimizing the expression and overall yield of N-CaM, the addition of calcium to expression media and during purification is important to producing and maintaining N-CaM in stabilizing conditions. Furthermore, this study demonstrates how it is useful to consider a ligand protectant during the production of proteins when a ligand is known to promote stability of the protein of interest.

2.2 Materials and Methods

2.2.1 Cloning and Construction of the Expression Plasmid

The N-terminal domain of calmodulin was obtained from the full length cDNA of calmodulin which was subcloned into the pET-15b plasmid described by previous methods (Urbauer and Topp laboratories) (*34*). Although this encodes the gene from chicken CaM, human

CaM has an identical amino acid sequence (*20*). The gene encoding N-CaM was obtained by using the PCR-based QuickChange method to mutate K77 to a stop codon in the cDNA of full length CaM. The following primers were used to generate N-CaM coding region: 5'-gatggcaagaaaaatgtaagatacagatagcgag-3' and 5'-ctcgctatctgtatcttacatttttcttgccatc-3'. PCR products were digested at 37 °C with Dpn I (Promega) for 1.5 hours. Immediately after digestion the DNA was transformed into XL-1 Blue Competent cells (Stratagene) and spread on LB agar plates containing 100 mg/mL ampicillin (LB/amp plates) then incubated at 37 °C overnight. Colonies isolated from the plates were cultured overnight at 37 °C in LB and ampicillin. The DNA from these cultures was isolated and purified using Qiagen Spin Miniprep Kit (Stratagene). The DNA containing CaM with a stop codon (taa) after M76 in the pET-15 vector was confirmed by sequencing the DNA using the T7 promoter and T7 terminator primers (Northwoods DNA, Inc., Bemidjii, MN).

2.2.2 Expression of N-CaM in *E. coli* cells

The pET-15 plasmid containing the gene for N-CaM (N-CaM-pET-15) was transformed into BL21 (DE3) cells (New England Biosciences, Ipswich, MA) and plated on LB/amp plates in which colonies were grown overnight at 37 °C. Starter cultures were made by inoculation with a colony into 10 mL of LB/amp cultures and grown overnight at 37 °C and shaken at 250 rpm. The starter cultures were added to 1 L of minimal media at pH 7.4 and grown further. An optical density (OD) reading is used as a measurement of the turbidity of the media at 600 nm, and in this case, the cells were grown to an OD_{600} of 0.6-0.8. Once cells reached this OD measurement, cells were then induced with a final concentration of 1 mM of isopropyl β-D-1-thioglactopyranoside (IPTG) (Acros Organics). Bacterial cells were harvested after 4 hours of

additional growth by centrifugation at 4,500 x g, and cell pellets were stored at -80 °C. To determine expression of N-CaM in the cultures, 1 mL samples were taken from each culture and boiled in 4 M urea and run using non-reducing 16% Tris-tricine SDS-PAGE for analysis by electrophoresis (*35*).

2.2.3 Purification of N-CaM

Cell pellets containing the overexpressed N-CaM were resuspended in 250 mM MOPS at pH 7.5, 500 mM KCl, 1 mM DTT, 1 mM EDTA and lysed by a French Press at 12,000 psi. Upon centrifuging the cells, a final concentration of 10 mM $CaCl_2$ was added to the soluble lysate containing N-CaM. The soluble lysate was injected into a fast-performance liquid chromatogrphy (FPLC) Äkta Explorer System (GE Healthcare) and applied to a phenyl sepharose column (GE Healthcare) equilibrated with 50 mM Tris-HCl pH 7.4, and 10 mM $CaCl_2$. The column was washed with 50 mM Tris-HCl pH 7.4, 10 mM $CaCl_2$, and 500 mM NaCl, and N-CaM was eluted with 50 mM Tris-HCl pH 7.4, and 10 mM EDTA. The fractions containing N-CaM were concentrated by ultrafiltration (Millipore). Because N-CaM possesses no tryptophans or tyrosines, the absorbance of phenylalanine at 259 nm is used to quantify N-CaM. The final yield was quantified by UV spectroscopy using the absorbance at 259 nm and an extinction coefficient of 975 M^{-1} cm^{-1}. The correct size and purity were identified using SDS-PAGE gels (*36*). Circular dichroism (CD) spectroscopy was used to determine N-CaM contained the expected secondary structural elements, indicating it was correctly folded. NMR analysis of N-CaM portrayed correct backbone assignments based on previously published assignments (*22*).

2.2.4 Screening of Ligand Concentrations in Expression Media

To obtain a high-yield of pure, intact N-CaM, modifications were made to the expression media and purification protocol. To improve the initial yield from the over expression of N-CaM, 6 different Ca^{2+} concentrations were screened in the expression media upon induction to the N-CaM-pET-15 plasmid. The flasks contained 20 mL of minimal media, prepared with 0.5 mM of $CaCl_2$ (before induction). A starter culture containing the N-CaM-pET15 plasmid in BL21 (DE3) cells were grown overnight and added to the 20 mL cultures to obtain a 100-fold dilution. The cells were grown at 37 °C and shaken at 250 rpm and induced when OD_{600} = 0.5 – 0.6. To find the optimal Ca^{2+} concentration that would result in an increase in the expression of N-CaM, $CaCl_2$ was added to the expression media upon induction. The appropriate volume of $CaCl_2$ was added to obtain the desired final Ca^{2+} concentration needed in each flask. The five different supplemental Ca^{2+} concentrations screened in the expression media were 1.0, 1.5, 2.0, 2.5, and 3 mM, and these were compared to one with no additional $CaCl_2$ added upon induction.

Size-exclusion chromatography was employed as a second purification step to remove higher molecular weight species that co-elute with apo N-CaM from the phenyl sepharose column and to exchange N-CaM into a final buffer system containing 100 mM $CaCl_2$ at low pH. The eluted apo N-CaM was immediately loaded onto a Superdex 75 10/300 GL column (GE Healthcare) equilibrated in 20 mM Mes, pH 5.5 and 100 mM $CaCl_2$. The chromatographic separation was monitored by a UV-Vis lamp at 220 nm and the final, pure and intact N-CaM eluted into 5 mL fractions. The fractions containing N-CaM were identified by 16% Tris-Tricine gels, then concentrated and quantified by UV-Vis spectroscopy. Final samples were stored in the equilibration buffer at 4 °C.

2.2.5 SDS-PAGE analysis and densitometry

To compare the levels of expression from each culture, samples were taken from each culture in 1-hour intervals over the course of 4 hours. The volumes were normalized to adjust for differences in cell growth rates, and the appropriate volume to obtain an equal number of cells was centrifuged at 13,000 rpm. The insoluble pellet was resuspended in saturated urea and Coomassie loading dye, then boiled at 95 °C for 10 minutes. Equivalent volumes were loaded onto 16% Tris-Tricine SDS-PAGE gels. After electrophoresis, the gels were stained in Coomassie dye and imaged by a Typhoon TRIO Variable Mode Imager from GE Healthsciences (Sweden). To compare the level of expression between cultures after 4 hours of growth, densitometry was used to obtain the relative intensities (reported as a volume-to-peak ratio) of the bands relative to the band intensity from the expression culture containing 0.5 mM $CaCl_2$. Each N-CaM band was normalized first by comparison to the intensity of a band from a 1 mg/ml of purified N-CaM on an SDS-PAGE gel.

2.3 Results

2.3.1 Influence of Ca^{2+} concentration of N-CaM expression in *E. coli*

The ability of a ligand to provide *in vivo* stabilization to facilitate the expression of N-CaM in bacterial cells was evaluated by screening a range of calcium concentrations in the expression media. Normally, the amount of calcium added to the expression media is dependent on the amount needed for the cells to maintain basal functions and to proliferate in minimal media. In this case, the amount needed for expression of N-CaM in minimal media was 0.5 mM $CaCl_2$. Figure 2.2 represents the expression of N-CaM in minimal media and 0.5 mM $CaCl_2$ over

the course of 4 hours. Maximum expression was observed at 4 hours, based on the intensity determined by densitometry.

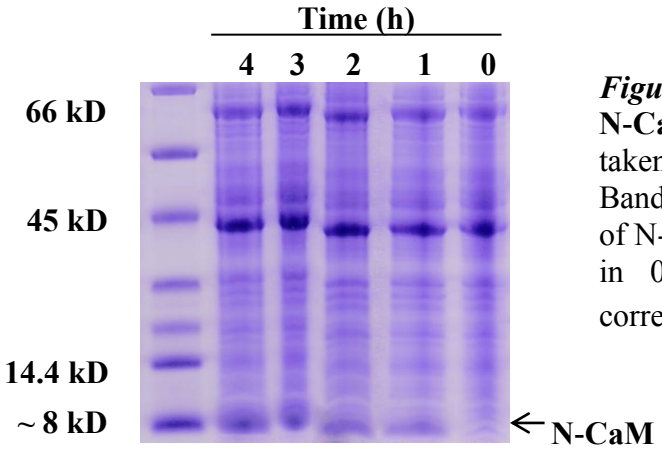

***Figure 2.2* SDS-PAGE gel of expression of N-CaM in *E coli* cells.** Gel samples were taken at each hour after induction (Time = 0 h). Bands at ~8 kDa corresponds to the expression of N-CaM (~8.5 kDa) over a 4 hour expression in 0.5 mM $CaCl_2$. 1st lane on the left corresponds to the molecular weight standards.

To determine the optimal amount of calcium needed to improve N-CaM expression, the additional amounts of calcium tested in this study were added immediately after the inducer, IPTG, was added to the media. A growth curve was generated for each of the cultures containing different calcium concentrations (Figure 2.3). The turbidity was measured at 600 nm and plotted as a function of time. The turbidity of the solution is an indication of the cell density in the cultures over the course of the cell growth.

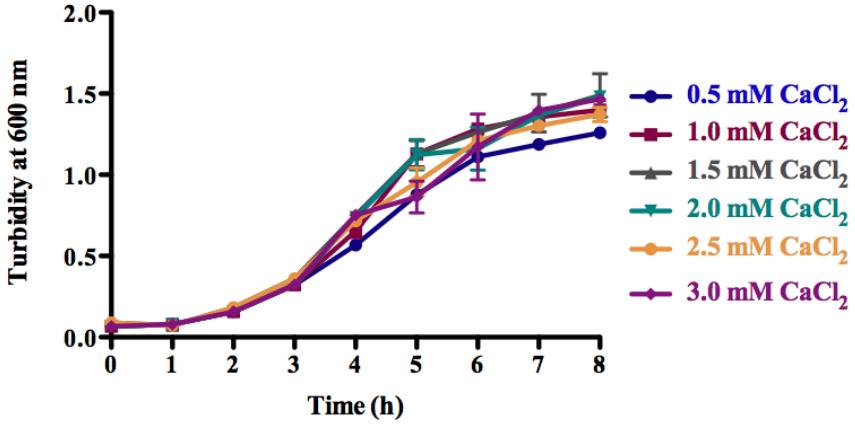

***Figure 2.3* Turbidity measurements of cultures.** Comparison of the cell densities from cultures with varying calcium concentrations. Time of induction was at 4 hours and cells were harvested at after 8 hours of growth.

Each curve displays a similar sigmoidal shape which can be separated into 3 phases: lag, exponential, and stationary phase. The cultures originated from the same colony grown in an overnight culture, which was diluted and then separated into aliquots so as to maintain similar growth profiles and induction periods for the test cultures. As such, the lag phases do not diverge substantially from one another. When comparing the exponential phase of each curve, any diversion of a curve compared to 0.5 mM $CaCl_2$ indicates the cell growth is altered slightly by the addition of calcium in the media. It is clear that at four hours after induction, the cell density is increased ultimately by the addition of calcium compared to no calcium, but not to a high degree. Because all curves have a similar shape and no large deviations occur, the cell population's ability to grow is not affected.

At 4 hours, samples from each of the cultures were taken and analyzed on SDS-PAGE gels to compare the expression levels of N-CaM (Figure 2.4). The bands present at ~8 kD, indicate N-CaM was expressed and present in all the cultures (data not shown). The relative intensities of the bands from each of the cultures were compared to the 0.5 mM $CaCl_2$ standard culture to determine the extent to which the additional calcium influenced the expression of N-CaM. The bands corresponding to the cultures containing 1.0, 1.5, and 2.5 mM $CaCl_2$ have similar intensities and each band appears to be darker than the band from 0.5 mM $CaCl_2$.

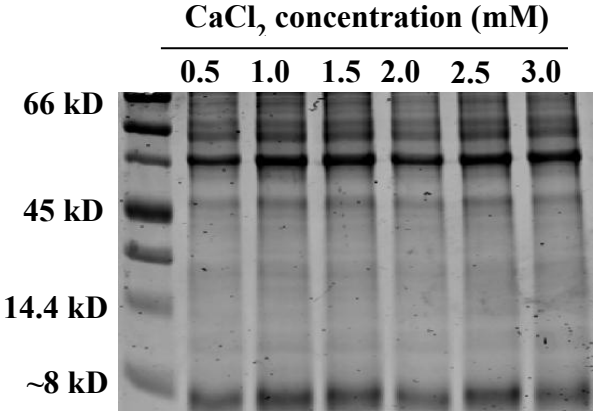

Figure 2.4 **SDS-PAGE gel of cultures with varying $CaCl_2$ concentrations.** Gel samples after 4 hours taken from each culture display bands at 8 kD corresponding to N-CaM. Bands were identified by comparing to purified N-CaM.

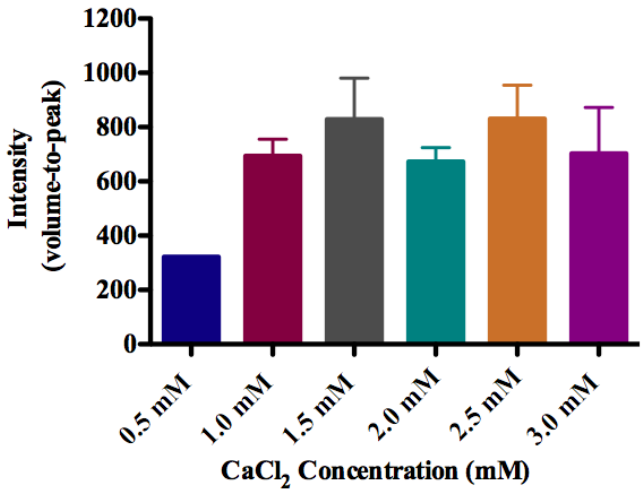

Figure 2.5 **Band intensities from cultures with varying $CaCl_2$ concentrations.** Comparison of the band intensities from gel samples at 4 hours of N-CaM expression in cultures with 0.5 mM – 3.0 mM $CaCl_2$ (volumes and peaks from SDS-PAGE bands determined by densitometry). Intensity values from samples 1.5 mM and 2.5 mM, were not significantly different based on a t-test. ($-2.92 < t < +2.92$, $t = 0.99$, $\alpha = 0.1$).

To determine the amount of calcium that is the most effective at increasing the expression of N-CaM, the intensities of these bands were compared using densitometry. The volume-to-peak intensities based on densitometric analysis of the bands from the culture with 1.0 – 3.0 mM $CaCl_2$ were compared against the culture containing 0.5 mM $CaCl_2$ (Figure 2.5). There is a clear increase in the intensity of the band from the expression of N-CaM when supplemental calcium is added, based on the increase in intensity compared to the culture with no additional calcium (0.5 mM $CaCl_2$). The cultures containing 1.5 and 2.5 mM $CaCl_2$, each display approximately a 2.5-fold increase in intensity. Because these intensities are not distinctly different it is difficult to discern whether the same amount of N-CaM is expressed in each of these two calcium concentrations without performing an overexpression and quantifying the final yield after purification of the crude lysate. The growth of the cells in 1.5 and 2.5 mM $CaCl_2$, however, are not affected based on the similar profiles compared to 0.5 mM $CaCl_2$ generated from the turbidity readings.

2.3.2 Effect of Ca^{2+} concentration on final protein yield

Expression of N-CaM in media containing 0.5, 1.5, and 2.5 mM $CaCl_2$ was performed to determine if higher expression of N-CaM results in improved yield of soluble, intact N-CaM, after purification. A two-step purification was performed on the samples to purify the N-CaM product, SDS-PAGE/densitometry was performed to determine the purity, and UV-Vis absorbance was used to quantify the final yield. Generally, CaM is purified employing hydrophobic-interaction chromatography (HIC) and exchanged into buffer containing calcium by dialysis. In this study, size-exclusion chromatography (SEC) was used to immediately reintroduce calcium and convert N-CaM from the apo to the holo state during the polishing step of the purification. N-CaM elutes from the HIC column in the presence of EDTA because release of Ca^{2+} from N-CaM decrease its surface hydrophobicity. Excess calcium promotes the binding of Ca^{2+} ions, which stabilizes N-CaM against chemical and physical degradation. Based on this known protection mechanism, we hypothesized the altering the purification method would result in a higher yield of pure N-CaM.

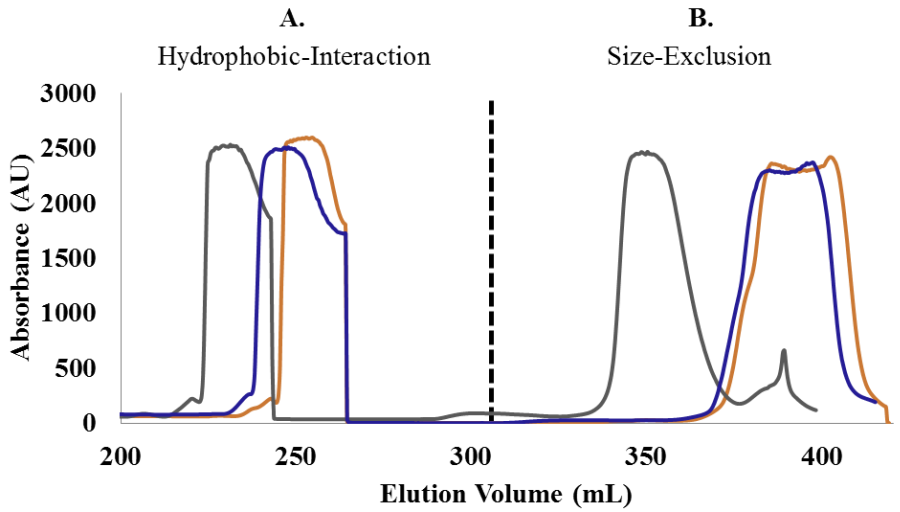

Figure 2.6 **Elution profiles comparing N-CaM purifications**. N-CaM expressed in 0.5 (blue), 1.5 (gray), and 2.5 mM $CaCl_2$ (orange) and the subsequent two-step purification. **A.** 1st peaks are the elution of apo N-CaM from a hydrophobic-interaction column **B.** 2nd peaks from the elution of holo N-CaM from a size-exclusion chromatography.

The resulting elution profiles from the two-step purification of each of the expressed N-CaM in different calcium concentrations are displayed in Figure 2.6. N-CaM was identified in the peaks displayed on Figure 2.6 and was determined by SDS-PAGE. The first peak (A) in the chromatogram corresponds to the elution of apo N-CaM from the hydrophobic interaction column. The second peak (B) corresponds to the elution of holo N-CaM from the size-exclusion column. Chromatography was employed to adequately separate N-CaM and isolate it from any contaminants in the solution. Figure 2.6 displays distinct, but broad peaks from each of these techniques to demonstrate N-CaM is effectively isolated from the bacterial lysate. The initial peaks resulting from the different purifications on the HIC column, display similarly shaped peaks. The shape of the peaks from 0.5 and 2.5 mM $CaCl_2$ from elution of the SEC column appear to be different from 1.0 mM $CaCl_2$ expression. The change in peak shapes may be due to the small change in the hydrodynamic radius as a consequence of the different expression media solutions. This was not investigated further because the correctly folded N-CaM was identified upon elution from the SEC column by NMR (see chapter 3).

Table 2.1 Elution volumes of N-CaM.

$CaCl_2$	V_e (mL)	
	HIC	**SEC**
0.5 mM	254	384
1.5 mM	231	350
2.5 mM	248	382

There are noticeable differences in the elution volumes (V_e) between each of the samples (Table 2.1). The peaks corresponding to 0.5 and 2.5 mM $CaCl_2$ expressions have similar elution volumes in the both chromatograms, while 1.5 mM $CaCl_2$ sample elutes earlier from each

column. Additionally, there are 2 peaks that appear in the SEC chromatogram for 1.5 mM $CaCl_2$ sample, in which the small peak appears at the same elution time as the 0.5 and 2.5 mM $CaCl_2$ samples. The SDS-PAGE analysis of the elution fractions revealed that every peak in the chromatogram contains N-CaM. Because there is limited information provided by SEC, it is difficult to determine the structural alterations that would result in a change in mobility through the SEC column. It is certain from SDS-PAGE data, however, that all peaks contain N-CaM and no traces of other protein contaminants.

The final yields of the N-CaM expressed in the different calcium concentrations and subsequent purification were compared by employing UV-Vis absorbance to quantify the final amount of N-CaM obtained after purification (Table 2.2). A substantial amount of protein was obtained from the expressions in 0.5, 1.5, and 2.5 mM $CaCl_2$. Compared to 0.5 mM $CaCl_2$, the supplemental amount of calcium in the expression media leads to a higher amount of N-CaM recovered after purification. Overall, purification of the sample with 1.5 mM $CaCl_2$ results in a 1.5-fold increase in the production of N-CaM per liter of media. Based on the turbidity readings and densitometric analysis of the expression of N-CaM, the addition of 1.5 mM $CaCl_2$ does not interfere with the normal cell growth or the ability of the cells to express N-CaM. From this data, 1.5 mM $CaCl_2$ was found to be an optimal amount of calcium needed to provide *in vivo* and *in vitro* stability to facilitate the expression and purification of N-CaM.

Table 2.2 Final yields of N-CaM from 1-liter cultures.

$CaCl_2$	N-CaM (mg)	SE[a]
0.5 mM	41	1.4
1.5 mM	62	2.1
2.5 mM	53	0.7

[a] standard error, where n = 2

2.4 Discussion

It is well known that calmodulin's ability to bind Ca^{2+} ions plays a role in its biological activity. Just as important, the binding of calcium can enhance the stability of N-CaM in various physiological and solution conditions (*24, 32*). In this study, the concept of using a ligand protectant to enhance stability was applied to developing a method to increasing the *in vivo* and *in vitro* production of pure, intact N-CaM. By using common approaches for expressing and purifying CaM, modifications were made to the calcium concentration in the expression media. In addition to the traditional methodology of applying hydrophobic-interaction chromatography for purification, a second purification step was involved to produce purified N-CaM in stabilizing solution conditions.

Upon evaluation of the expression of N-CaM in varying amounts of calcium, SDS-PAGE gels and turbidity measurements revealed that additional calcium had significant impact on the amount of N-CaM expressed and a no large effect on the growth of the cells. Based on the increase of turbidity of the solution, adding more than 0.5 mM $CaCl_2$ to the expression media alters the overall cell density slightly. An increase in cell density was observed when calcium was added to the expression media 4 hours after the cell growth was initially started (the cells grow for 4 hours and calcium is added at induction). Calcium regulation in the cell is an energy-intensive process relying on various types of cellular machinery. A balance is maintained in cells to keep intracellular levels of Ca < 0.1 mM and extracellular levels ≤ 1 mM depending on the type of cell (*37*). The alteration of the nutrients in the environment of the bacterial population may have an impact on the metabolic processes of the bacteria, which may impact the growth rate. Whether the addition of calcium changes the growth rate of the population is not clear. Monitoring the cell density at 600 nm measures the turbidity of the solution and not the actual

amount of cells in solution (cannot discern between alive and dead cells) over a period of time. Precipitation of additional calcium in the media could also affect the optical density, but there is no way to easily account for this. Based on the consistency of the curve shape from all turbidity readings, cell growth is not substantially altered. The premise for adding excess calcium to the expression media upon adding an inducer is to have calcium readily available upon expression of N-CaM, but not to affect the ability of the cells to grow or generally impair the function of the protein expression machinery. For this reason, low calcium concentrations were added at the time N-CaM is being expressed and not prior to expression. Having available free Ca^{2+} ions in the media and cytosol at the time N-CaM is released could facilitate the coordination between Ca^{2+} ions and residues in binding sites I and II of N-CaM.

Many physiological and environmental factors can influence the degradation of proteins *in vivo* and contribute to a low expression of protein. In *E. coli*, proteins expressed in the cytosol (soluble proteins) are susceptible to chemical and physical degradation, which may result in their destruction by cellular proteases. These degradation pathways are influenced by the physiological and external environments, such as temperature and pH. N-CaM is expressed as a soluble protein and the susceptibility of the known degradation sites contained in N-CaM can be influenced by these environmental stresses.

Upon expression and folding in the cytosol, exposed degradation sites in N-CaM are susceptible to chemical modifications via deamidation of the sidechain of asparagine 60 (N60), racemization of aspartates 22, 24, and 58, and oxidation of methionine residues 51, 71, and 72 (*9, 38*). CaM has increased susceptibility to deamidation when not bound to Ca^{2+} ions, and deamidation occurring to Asn sidechains in the Ca^{2+}-binding sites are more likely to undergo destruction by the organism's proteasome (*39*). Methionine residues become susceptible to

oxidation upon binding Ca^{2+} in the presence of oxidants (*i.e*. hydrogen peroxide). Oxidation of methionine residues 51, 71, and 72 can assist in the unfolding of CaM and has known to be a signal for destruction by the 20S proteasome (*38*). Studies inducing oxidation of Met residues were at high concentration of hydrogen peroxide and our expression or purification conditions do not expose N-CaM to such oxidizing conditions. Expression conditions, however, expose N-CaM to high temperatures at pH 7 for ~4 hours and could promote deamidation and/or isomerization promoting increased accumulation of N-CaM. Having available free Ca^{2+} ions upon expression of N-CaM is most likely protecting it from deamidation and/or isomerization. Ligand binding, in this case, is limiting the flexibility of labile residues in the binding sites and reducing conformational exchange, thereby protecting N-CaM from cellular degradation.

Quantification of the band intensities from the expressions, demonstrates that increasing the calcium concentration to a point improves the overall expression of N-CaM. The higher intensities of bands from the expressions in 1.5 and 2.5 mM $CaCl_2$, yet slightly higher yield of protein at 1.5 mM $CaCl_2$ indicate that not all calcium concentrations affect the final yield exactly the same. The calcium concentrations 1.5 and 2.5 mM $CaCl_2$ seem to follow the same trend as the other concentrations by altering the cell density slightly; a larger amount of protein is, however, obtained from 1.5 and 2.5 mM $CaCl_2$. Based on the actual amount obtained between these two samples, 1.5 mM $CaCl_2$ in the expression media yields a higher amount of pure protein compared to 0.5 and 2.5 mM $CaCl_2$.

It is likely during any purification process that low recovery of protein can be attributed to physical and/or chemical instability. In the case of N-CaM, low recovery could be due to the different stages where N-CaM is not bound to calcium during the two-step purification process. Because the apo form of N-CaM is more susceptible to physical and chemical degradation, the

resulting degradation products could lead to a lower recovery of protein. At different stages, the addition and chelation of Ca^{2+} ions in solution containing N-CaM facilitates conformational exchange between the apo and holo forms of N-CaM. Upon binding to calcium, N-CaM exposes hydrophobic patches known to bind to peptide targets (*20-22, 30*). It is possible that exposure of hydrophobic residues in the absence of these targets could promote intermolecular interactions leading to self-associated aggregates. Before the purification process, precipitation is observed after adding calcium to the crude lysate containing N-CaM. Centrifugation is required to separate precipitates from soluble N-CaM. The remaining soluble N-CaM undergoes purification, and no precipitation is seen thereafter. Based on SDS-PAGE analysis of the precipitation, N-CaM is observed in the insoluble lysate. Precipitation is an indication of physical degradation of protein and is often associated with insoluble aggregate formation (*3*). Precipitation is normally seen at this step and makes no difference in which calcium conditions N-CaM was expressed. Based on the final yield of N-CaM after purification, 1.5 mM $CaCl_2$ improves the final recovery of protein overall.

2.5 Conclusion

Based on the final recovery of protein, it seems there is a balance that must be maintained between providing enough supplemental calcium to stabilize N-CaM, thereby promoting high-yield expression, and excess calcium that will diminish overall yield. Because 1.5 mM $CaCl_2$ improves the initial expression and yields the highest amount of purified protein, it is the best amount of calcium to supplement the expression media. Additionally, 1.5 mM $CaCl_2$ is an adequate amount to overcome physical and chemical degradation processes that could result in low recovery. The final yield of pure, intact N-CaM expressed in 1.5 mM $CaCl_2$ provides an

increase in protein yield of 50%. This is a sufficient amount of protein for analysis by NMR spectroscopy, which is described in the next chapter.

2.6 References

1. Biekofsky, R.R., et al., *Ca2+ coordination to backbone carbonyl oxygen atoms in calmodulin and other EF-hand proteins: 15N chemical shifts as probes for monitoring individual-site Ca2+ coordination.* Biochemistry, 1998. **37**(20): p. 7617-29.
2. Finn, B.E. and S. Forsen, *The evolving model of calmodulin structure, function and activation.* Structure, 1995. **3**(1): p. 7-11.
3. Ikura, M., *Calcium binding and conformational response in EF-hand proteins.* Trends Biochem Sci, 1996. **21**(1): p. 14-7.
4. Jurado, L.A., P.S. Chockalingam, and H.W. Jarrett, *Apocalmodulin.* Physiol Rev, 1999. **79**(3): p. 661-82.
5. Klee, C.B., T.H. Crouch, and P.G. Richman, *Calmodulin.* Annu Rev Biochem, 1980. **49**: p. 489-515.
6. Mal, T.K. and M. Ikura, *NMR Investigation of Calmodulin.* Modern Magnetic Resonance, 2006(Part 1, Part 6).
7. Masino, L., S.R. Martin, and P.M. Bayley, *Ligand binding and thermodynamic stability of a multidomain protein, calmodulin.* Protein Sci, 2000. **9**(8): p. 1519-29.
8. Yang, J.J., A. Gawthrop, and Y. Ye, *Obtaining site-specific calcium-binding affinities of calmodulin.* Protein Pept Lett, 2003. **10**(4): p. 331-45.
9. Sorensen, B.R. and M.A. Shea, *Interactions between domains of apo calmodulin alter calcium binding and stability.* Biochemistry, 1998. **37**(12): p. 4244-53.
10. Weickert, M.J., et al., *Optimization of heterologous protein production in Escherichia coli.* Curr Opin Biotechnol, 1996. **7**(5): p. 494-9.
11. Bartlett, R.K., et al., *Oxidation of Met144 and Met145 in calmodulin blocks calmodulin dependent activation of the plasma membrane Ca-ATPase.* Biochemistry 2003(42): p. 9.
12. Schägger, H., *Tricine-SDS-PAGE.* Nature Protocols, 2006(1): p. 6.
13. Masino, L., S.R. Martin, and P.M. Bayley, *Ligand Binding and thermodynamic stability of a multidomain protein, calmodulin.* Protein Science, 2000. **9** p. 10.
14. Balog, E.M., et al., *Site-specific methionine oxidation initiates calmodulin degradation by the 20S proteasome.* Biochemistry, 2009. **48**(13): p. 3005-16.
15. Potter, S.M., W.J. Henzel, and D.W. Aswad, *In vitro aging of calmodulin generates isoaspartate at multiple Asn-Gly and Asp-Gly sites in calcium-binding domains II, III, and IV.* Protein Sci, 1993. **2**(10): p. 1648-63.
16. Tarcsa, E., et al., *Ca2+-free calmodulin and calmodulin damaged by in vitro aging are selectively degraded by 26 S proteasomes without ubiquitination.* J Biol Chem, 2000. **275**(27): p. 20295-301.
17. Wang, W., *Instability, stabilization, and formulation of liquid protein pharmaceuticals.* Int J Pharm, 1999. **185**(2): p. 129-88.

Page Left Intentionally Blank

CHAPTER 3

USING NMR SPECTROSCOPY TO INVESTIGATE HOW SOLUTION CONDITIONS INFLUENCE PROTEIN STABILITY

3.1 Introduction

Many studies involved in developing methods to study protein stability use calmodulin (CaM) as a model system because the thermal stability of the whole, intact CaM and the isolated N- and C-terminal domains are well-characterized. Studies that contribute to the vast amount of literature in which the physical stability of CaM is characterized typically use low-resolution spectroscopic methods to examine stabilization (*9, 23, 24, 32*). Structural and thermodynamic information is obtained using fluorescence and absorbance-based methods to determine the influence of calcium and different solution conditions on global structure and/or stability. Although these methods provide a rapid approach to analysis of thermal stability, site-specific molecular information is not acquired from these methods. When combined with such general structural and thermal stability information, molecular information provides essential details for understanding the mechanism of overall stabilization. Residue-specific information required for characterizing stability on a molecular level may be obtained through NMR experiments.

3.1.1 Structural and Thermal Stability of CaM

As one of the major calcium sensors in many organisms, CaM belongs to the EF hand family of calcium-binding proteins. The EF hand within these proteins is a helix-loop-helix motif in which alternating residues within the loop region coordinate to one Ca^{2+} ion. The well-known dumbbell structure of CaM is composed of two homologous globular domains separated by a

short and flexible antiparallel β-sheet linker (Figure 2.1). The N- and C-domains each consists of a pair of EF hands, resulting in a C2 axis of symmetry per domain. Figure 2.1 displays the two known conformers of CaM, where the apo form is structurally compact compared to the extended holo form and CaM, upon binding to calcium, increases in helicity based on circular dichroism analysis (*21*).

The four Ca^{2+}-binding loops are similar in sequence, each possessing 12 residues in the binding site, which are highly conserved amongst the EF hand family of proteins. Six of these residues are involved in calcium coordination, in which Ca^{2+} is coordinated in a pentagonal bipyramidal orientation (Figure 3.1) (*31*). Because each domain is composed of a pair of EF hands, the binding sites are also structurally equivalent. One major difference is the N-domain has lower affinity for Ca^{2+} compared to the C-domain ($K_d \sim 10^{-7}$). The highly conserved sequence in CaM amongst organisms displays similar biochemical and physical properties as demonstrated by many studies that utilize spectroscopic methods(*20*).

The N- and C-domains of CaM are homologous to each other with 46% sequence identity (*40*). As such, the two domains have similar NMR spectra, and for simplicity, this study uses the N-domain of CaM (N-CaM) as a model system for development of a method to utilize two-dimensional (2D) NMR spectroscopy to study protein stability. This truncated form of CaM, is known to remain functional and well-structured as an isolated domain. Previous studies have determined the thermal stability of the N-terminal domain within the intact protein and as an isolated domain and have evaluated how the overall thermal stability and folding of the domain is impacted by calcium (*24*). Spectroscopic methods have been used to show that calcium binding and ionic strength each impart thermal stabilization to full-length CaM and the isolated domains (*24*).

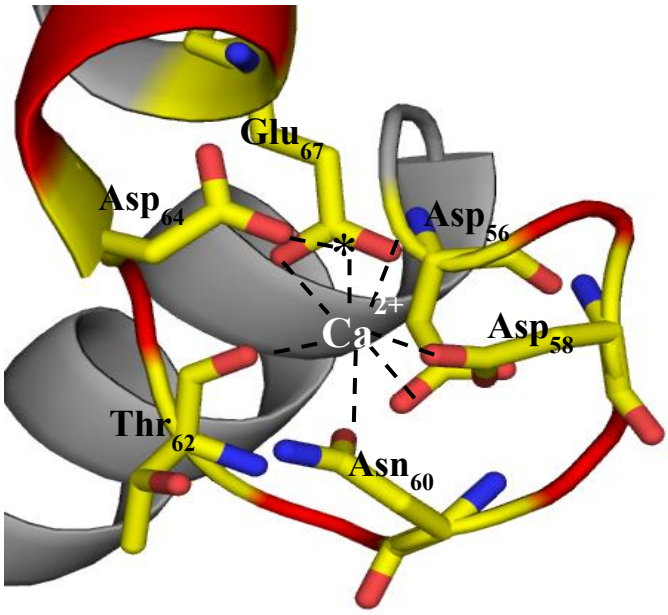

Figure 3.1 **Ca^{2+}-binding site II in the N-terminal domain of CaM.** The binding loop in binding sites I-IV is structurally equivalent and composed of highly conserved residues at positions 1, 6, 12, which in site II are Asp 56, Gly 61, and Glu 67, respectively. Structure from PDB ID 3cln.

3.1.2 Developing a method to study protein stability by NMR spectroscopy

A two-dimensional (2D) ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiment was used in this study to evaluate residue-specific changes in apo N-CaM in a range of solution conditions known to affect the protein's thermal and structural stability. 2D NMR spectroscopy is a valuable tool to study changes in conformation, dynamics/flexibility and to assess local structure. To monitor local structure by 2D NMR, many studies have used ^{1}H-^{15}N HSQC experiments to selectively detect chemical shifts of individual amide moieties from residues of interest. Data from ^{1}H-^{15}N HSQC experiments contain site-specific information to facilitate the determination of residues that coordinate to Ca^{2+} and to assess local conformational changes that occur upon Ca^{2+} coordination and thermal perturbation (*19, 22*).

The ^{1}H-^{15}N HSQC experiment detects changes in structure and is sensitive enough to detect subtle conformational changes that occur locally at each backbone amide upon modest alteration of the solution conditions. This is because the electronic environments of individual

nuclei are affected differently by changing solution conditions. In a 2D ^{1}H-^{15}N HSQC spectrum of a protein a cross-peak corresponds to a proton covalently bonded to a nitrogen atom, or an amide group within the protein. Based on the degree of change in the position of the peak, qualitative information can be obtained to indicate how the change in environment physically affects an individual residue. Similarly, the changes in the dynamics/flexibility of the protein in solution may also be assessed by monitoring the broadening or sharpening of a peak that may occur as a consequence of changing the solution conditions (*6*). To evaluate structural stability, 2D NMR experiments can easily be used to screen a range of solution conditions and to monitor changes in chemical shift positions in the ^{1}H and ^{15}N dimensions. NMR provides a unique approach to analyze local structure and conformational changes and serves as a complimentary tool to other spectroscopic techniques to evaluate molecular mechanisms of physical stability.

3.2 Materials and Methods

3.2.1 Materials

Calcium chloride dihydrate was purchased from ACROS ORGANICS (Fair Lawn, NJ), potassium chloride (Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), (4-(2-hydroxyethyl)-1-piperazineethane-sulfonic acid (Hepes), and 3-(4-morpholino)propane sulfonic acid (Mops), 2-(4-morpholino)ethanesulfonic acid (Mes), ethylenediamine tetraacetic acid disodium salt dihydrate (EDTA) and sodium chloride were all purchased from Fisher Scientific (Pittsburg, PA). For NMR experiments, ^{15}N-ammonium chloride and deuterium oxide (D_2O) were purchased from Cambridge Isotopes (Andover, MA).

3.2.2 Expression and Purification of ^{15}N-labeled N-CaM

To obtain ^{15}N-labeled protein for NMR analysis, N-CaM was expressed and purified as previously described (Chapter 2) with the following modifications. The bacterial cells were grown in 1 L of minimal media containing [^{15}N]-labeled ammonium chloride. Upon induction, $CaCl_2$ was added to make the final calcium concentration 1.5 mM. Upon purification by size-exclusion chromatography, pure ^{15}N-labeled N-CaM was stored in 20 mM Mes, pH 5.5 and 100 mM $CaCl_2$ at 4 °C.

3.2.3 NMR Sample Preparation

Apo N-CaM was obtained by dialyzing 0.5 mL of holo N-CaM in 3.5k MWCO Slide-a-Lyzer dialysis cassettes composed of a 30 μm cellulose membrane (Thermo Scientific, PA) against 20 mM Mes, pH 5.5 and 100 mM EDTA overnight at room temperature (r.t.) to remove Ca^{2+} ions from N-CaM. Dialysis was performed in 1 L for approximately 12 hours. Upon the completion, apo N-CaM was removed from the dialysis cassettes and quantified by UV absorbance and was diluted into the same dialysis buffer with 6% D_2O to obtain a final protein concentration of 1.0 - 1.5 mM. A pH titration of apo N-CaM was performed at pH 5.5, 6.0, 6.4, or 6.7 by dialyzing 0.5 mL of N-CaM against 1 L of 20 mM Mes at the desired pH in 10 mM EDTA for 1 hour. To obtain apo N-CaM at pH 7.1 and 7.4, the protein was dialyzed into 20 mM Hepes and 10 mM EDTA.

To analyze the effect ionic strength has on the physical stabilization of apo N-CaM, an ionic strength titration was performed at pH 6.4 and 7.1. The ionic strength concentrations evaluated at each pH were 35, 65, 95, 125, and 155 mM. The total ionic strength of each solution

was determined by taking into account the contributing ionic species from the buffer and EDTA components at each pH. Equation 1 was used to determine the total ionic strength (*I*), in which M_i is the molar concentration of all ionic species and Z_i is the corresponding charge at each pH.

$$I = \sum_i M_i Z_i^2 \qquad (1)$$

To obtain apo N-CaM at either pH 6.4 or 7.1, the initial dialysis was performed for 2 hours in either 20 mM Mes, pH 6.4 or 20 mM Hepes, pH 7.1, 10 mM EDTA at *I* = 35 mM. Subsequent dialysis steps at increasing ionic strength were performed for 1 hour. Upon completion of each dialysis step, the sample was measured by using a pH meter (Accumet, Fisher Scientific) calibrated with 3-points at pH 4.00, 7.00, 10.00, to confirm the sample was at the expected pH.

3.2.4 NMR Experiments

The ^{1}H-^{15}N HSQC experiments were performed on either a Varian Inova 600 or a Bruker Avance 600 MHz spectrometer equipped with a triple resonance probe. The spectra were acquired in 8 scans with 2048 points in ^{1}H and 128* increments in ^{15}N. An external DSS standard in D_2O was used for referencing the chemical shift position of ^{1}H. Indirect referencing of ^{15}N to ^{1}H as determined by using the frequency ratio method ^{15}N/^{1}H = 0.101329118 (*41*). Sparky was used to determine positions, heights, and volumes of peaks in the spectra acquired from the ^{1}H-^{15}N HSQC experiments (*42*).

Assignments of the residues in N-CaM were based on Ikura *et al.(43),* and a pH titration was performed to correlate assignments at the reported pH with those obtained over the pH range 5.5 to 7.4 (*22*). The ^{1}H-^{15}N HSQC data were analyzed further to compare changes in peak positions from the resulting pH and ionic strength titrations of apo N-CaM at 25 °C. To analyze

the changes in peak positions (Δδ) in the spectra from the pH and the ionic strength titrations, changes in chemical shift (ppm) in the ^{1}H dimension for individual residues at each pH or ionic strength were quantified. This was calculated by subtracting the peak position at the higher pH or ionic strength (δ_x) from the initial peak position at the initial pH or ionic strength (δ_i), as given in Equation (2). Because of this approach to identify peaks, only unidirectional changes were monitored, and changes in the ^{15}N ppm were not included in the analysis.

$$\Delta\delta_x = \delta_x - \delta_i \qquad (2)$$

3.2.5 Circular Dichroism Sample Preparation

The thermal stability of apo N-CaM was monitored by circular dichroism (CD) at pH 5.5, 6.4, and 7.1 (I = 5 mM). Additionally, the thermal stability at pH 6.4 and 7.1 was analyzed over the ionic strength range of 5 – 125 mM to assess the impact of ionic strength on stabilization at higher pH. Apo N-CaM was was obtained by the same methods for NMR experiments, with the exception of using ^{14}N-ammonium chloride in the expression media. To obtain samples at pH 5.5 and 6.4, apo N-CaM was dialyzed into 10 mM Mes, 1 mM EDTA for 2 hours at r.t. To obtain pH 7.1, apo N-CaM was dialyzed into 10 mM Hepes and 1 mM EDTA for 2 hours at r.t. The buffer for each dialysis was adjusted by adding KCl to obtain the final desired ionic strength. Following dialysis, the protein concentration was quantified based on UV-Vis spectroscopy (absorbance at 259 nm, ε = 975 M^{-1} cm^{-1}) and adjusted to a final concentration of 0.2 mg/mL for subsequent analysis.

3.2.6 Thermal Stability of Apo N-CaM Evaluated by Far-UV CD

Far-UV circular dichroism measurements were acquired on an Olis MultiScan Spectrophotometer (On-line Instrument Systems, Inc, Borgart, GA; supported by Olis SpectralWorks software for data analysis) equipped with a Peltier temperature controller with a 6-position cell holder. To assess the helical secondary structure of apo N-CaM, the signal at 222 nm was monitored over the temperature range of 10 - 87.5 °C and spectra were obtained in 2.5 °C increments, utilizing a temperature ramp of 15 °C/hr. The CD data acquired from the thermal titrations, were converted into units of molar ellipticity and plotted as a function of temperature(*12*). Error analysis was performed from duplicate samples, and the data were fit to a sigmoidal function to obtain transition melting temperatures (T_m) using GraphPad Prism version 5.04 for Windows (Graphpad Software, San Diego, CA).

3.3 Results

3.3.1 The Influence of pH on the Structure of Apo N-CaM Assessed by NMR Spectroscopy

A spectrum of the purified N-CaM in pH 5.5 at 25 °C was determined to be the correct structure based on previous 2D NMR spectra and published assignments of apo CaM by Ikura *et al*. (Figure 3.2.A) (*22*). Residues were identified in each of the Ca^{2+}-binding loops of site I and II, adjoining beta strands, and outside of the binding sites based on previous resonance assignments of the residues in CaM (Figure 3.2.B) (*22*). A good indication that the apo form is well-structured is the appearance of the characteristic high-field amide protons (9.0-10.0 ppm) that correspond to the glycines at the sixth position of each of the two Ca^{2+}-binding loops. This structural and spectral feature of properly folded, functional Ca^{2+}-binding proteins is common among the EF hand family (*22*).

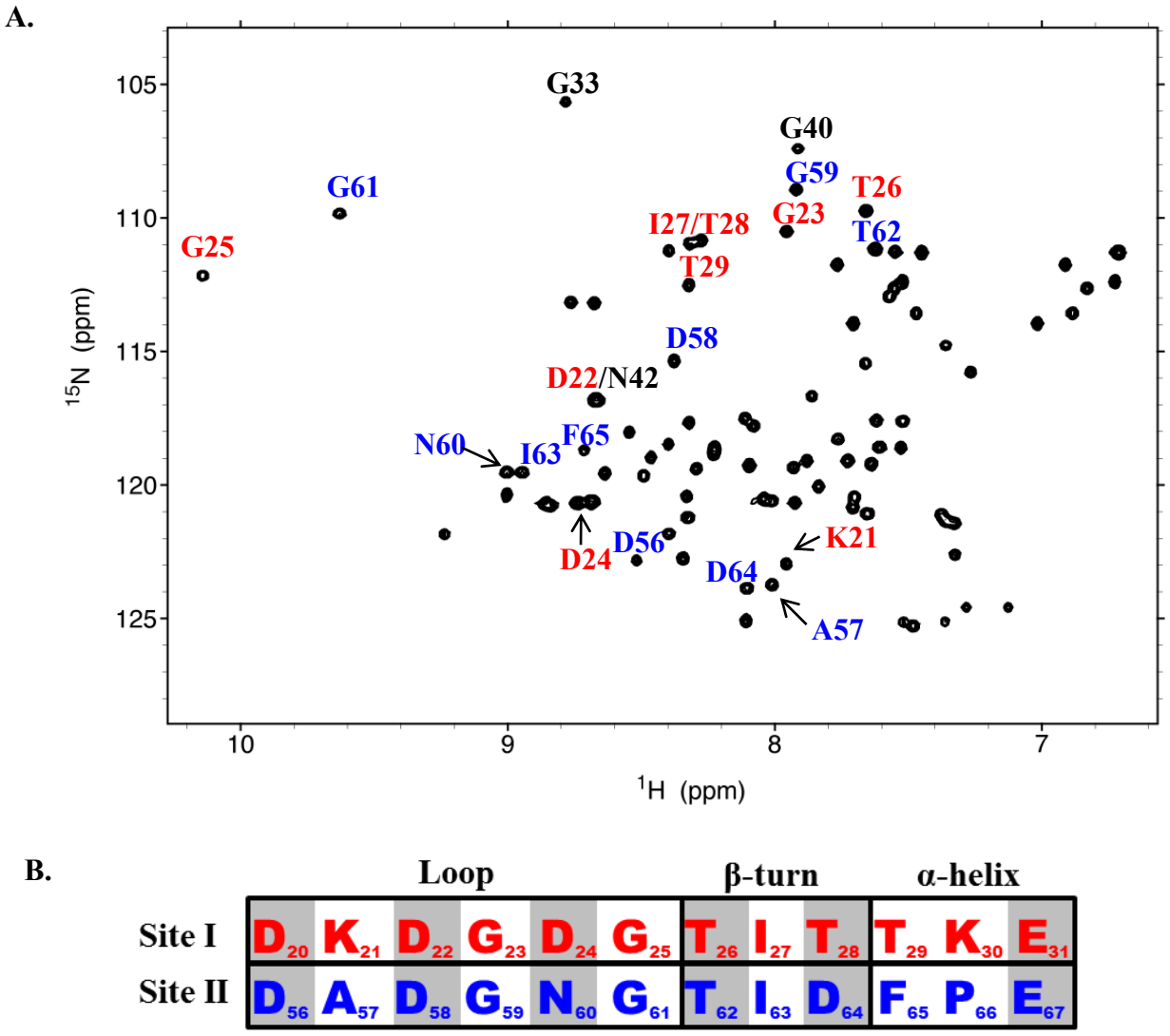

***Figure 3.2* A ^{1}H-^{15}N HSQC spectrum of apo N-CaM.** Apo N-CaM (1 mM) in 20 mM Mes, 10 mM EDTA at pH 5.5. Residues in Ca^{2+}-binding sites I (red) and II (blue) and residues outside of binding loops in α-helical region (black). Data acquired on a Varian Inova 600 MHz**. B. Residues in Ca^{2+}-binding sites of N-CaM.** Binding sites I (red) and II (blue) and corresponding secondary structural elements in the N-terminal domain of Calmodulin. Residues coordinated to Ca^{2+} ion are shaded in gray.

Some of the residues of interest could not be identified or tracked in our spectrum due to either peak overlap or extreme crowding, such as in the center of the spectrum. A pH titration of apo N-CaM was used to facilitate the identification of peak assignments of residues in N-CaM over a broad pH range. By tracking the unidirectional movement of individual peaks, chemical shift position can be correlated at each pH with the reported assignments (*22, 43*). Additionally, the pH titration provides evidence of structural perturbations at individual nuclei, thereby reporting local structure as a consequence of increasing the pH. Specifically, by monitoring the chemical shift positions of individual residues within the Ca^{2+}-binding sites I and II at each pH, spectra from ^{1}H-^{15}N HSQC experiments can provide information about local conformational changes and structural perturbations.

^{1}H-^{15}N HSQC experiments used to monitor apo N-CaM in pH 5.5 at 25 °C over the course of one week confirm that no large chemical shift changes are observed in apo N-CaM over this time frame (data not shown). Consequently, the pH titration of apo N-CaM was performed beginning at pH 5.5 to minimize chemical and physical degradation during storage and analysis. Following each ^{1}H-^{15}N HSQC experiment, the sample was dialyzed progressively to higher pH and the additional experiments were performed to yield a series of spectra over the pH range of 5.5 to 7.4 (Figure 3.3). Based on the NMR data and previously reported deamidation rates (*8, 9*), no more than 8% degradation occurs to N-CaM during the entire data acquisition period. Such a small amount of degradation produces only a very weak intensity NMR signal compared to intact N-CaM.

The spectra of apo N-CaM at different pH were compared and, by visually tracking the changes in peak position, qualitative analysis of the influence of pH on individual residues and overall structure is apparent. Figure 3.3 shows an overlay of spectra acquired at several pH

values during this titration. The majority of the resonances of individual residues from each spectrum closely overlap; this lack of change in chemical shift positions indicates that the global structure is not affected over the course of the titration. A few individual peaks are affected considerably by the obvious, large changes observed in the linear movement of peak position with increasing pH. The peaks circled in Figure 3.3 correspond to individual residues that undergo relatively large changes in peak position, which are N60, G61, and D64, located in the binding loop of site II. Based on this series of spectra, adjacent residues in binding site II are not altered by pH. These three residues (N60, G61, and D64) follow a common trend and move upfield substantially with increasing pH. This type of chemical shift movement indicates increasing pH has an effect on local structure such that the nuclei become more deshielded and deviate further from random coil values.

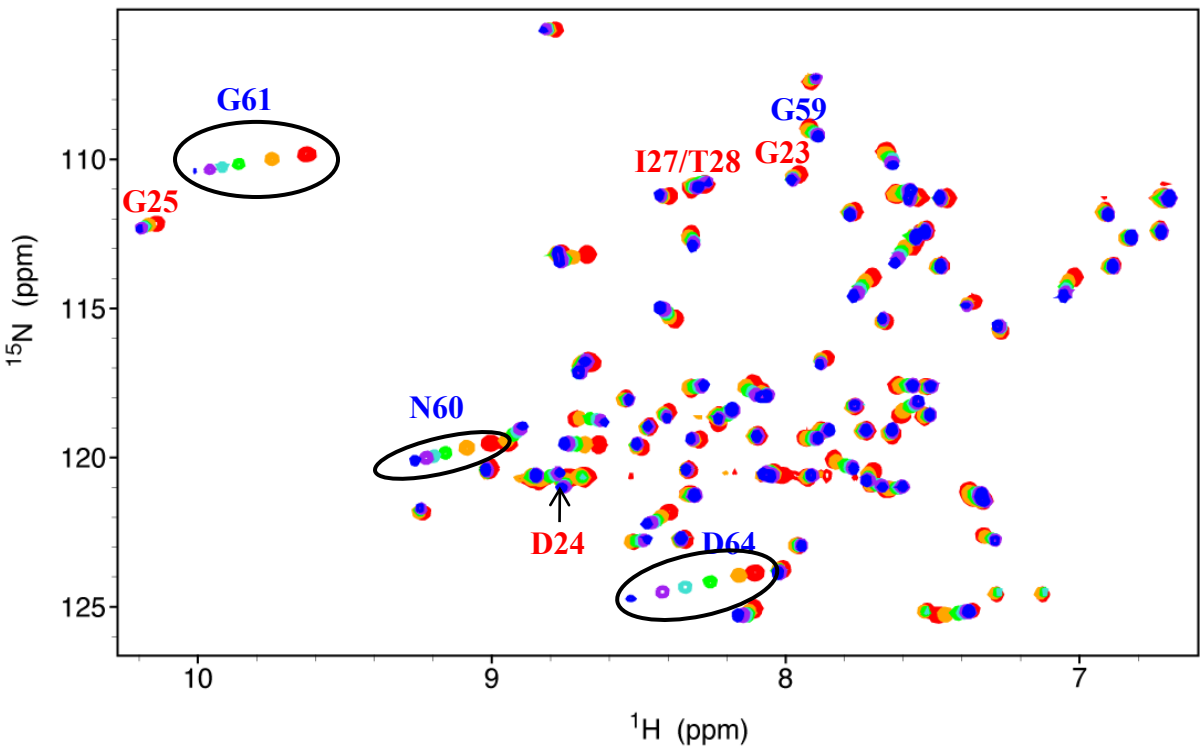

Figure 3.3 **A series of spectra of apo N-CaM from ^{1}H-^{15}N HSQC experiments in increasing pH.** Peaks are colored based on pH conditions: pH 5.5 (red), pH 6.0 (orange), pH 6.4 (green), pH 6.7 (turquoise), pH 7.1 (purple), pH 7.4 (blue). Residues in Ca^{2+}-binding sites I (red) and II (blue).

The residues in binding site II that displayed large deviations are in the same type of secondary structural elements as the residues in corresponding positions in binding site I (Figure 3.2.B). Because they are in equivalent structures, the relative change in chemical shift in the ^{1}H dimension (Δδ) of the residues in site I were compared to the Δδ of the residues in site II as a function of pH (Figures 3.4.A and 3.4.B). As depicted in the spectra and Figure 3.4.A, the residues in positions G23, D24, G25 (GDG sequence) in site I do not display large changes in chemical shift position. Residues in the GDG sequence undergo small movements in peak positions; deviations from their original peak positions are less than 0.056 ppm.

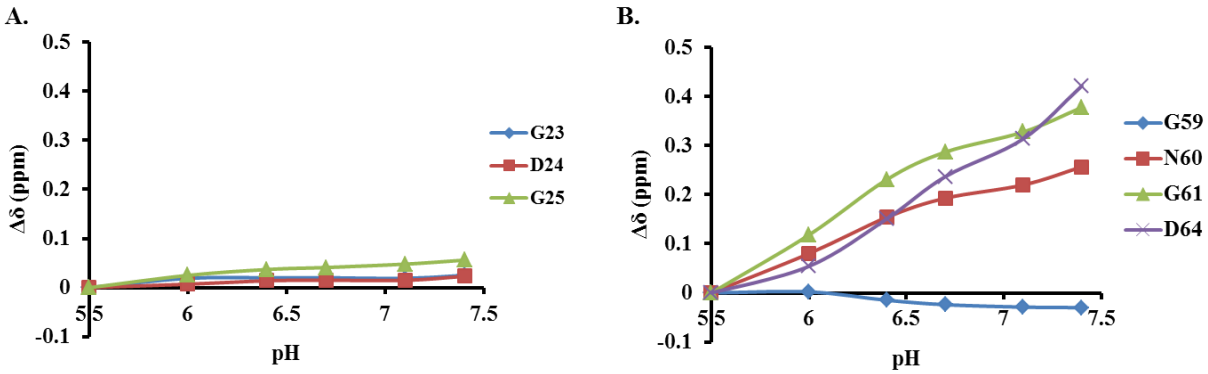

Figure 3.4 **Change in chemical shift position (Δδ) of residues from pH titration.** The Δδ of residues in Ca^{2+}-binding loops, based on ^{1}H-^{15}N HSQC spectra from the pH titration. **A.** Residues in Ca^{2+}-binding site I. **B.** Residues in Ca^{2+}-binding site II.

Figure 3.4.B displays residues G59, N60, G61 (GNG sequence), which are in the analogous positions to G23-G25 in the loop of binding site I. Additionally, D64 is also plotted in Figure 3.4.B, which is positioned in the short β-turn structure. Both the spectra and Δδ, clearly show that N60, G61, and D64 follow a similar trend in which the peaks move upfield in a linear trajectory with respect to chemical shift position. The Δδ for residues in binding site II is larger than site I and varies up to 0.422 ppm, confirming that local conformational changes are more

pronounced compared to the residues in binding site I. The G59 residue does not follow the same trend, but instead undergoes a relatively small downfield shift with increasing pH, like G23 in site I. Residue D64 experiences the largest deviation in original peak position compared to the GDG and GNG residues in binding site I and II, respectively. In the same position of binding site II, the corresponding peak in binding site I, residue T28, is partially overlapped with I27 and $\Delta\delta$ cannot be determined accurately. Nonetheless, it is clear from the spectra that T28 appears to be affected similarly to adjacent residues in binding site I and does not undergo obviously large shifts in peak position.

3.3.2 Ionic Strength Titration of Apo N-CaM Evaluated by NMR Spectroscopy

The ^{1}H-^{15}N HSQC experiments were performed to assess how the ionic strength affects individual residues in apo N-CaM. High ionic strength is known to enhance the thermal stability of CaM at pH 8, but the molecular mechanism of stabilization is not known (*24*). The present study tests the hypothesis that stabilization is achieved by affecting specific residues in apo N-CaM. By exposing apo N-CaM to high pH and conducting an ionic strength titration, residues which play a role in the stabilizing effect of KCl will be revealed. The effect of increasing ionic strength (*I*) on individual residues of apo N-CaM was assessed by comparing a series of ^{1}H - ^{15}N over the range of 35 – 155 mM (selected spectra shown in Figures 3.5 and 3.7).

The ionic strength titrations were performed at pH 6.4 and 7.1 and conducted under a sufficiently short time frame to minimize chemical degradation processes that occur at higher pH to less than 8%. The overlay of each spectrum from the ionic strength titration at pH 6.4 did not show significant changes in peak positions for the majority of residues or large peak position changes in any one residue (data not shown). To show the minimal effect that increasing ionic

strength had on the global fold of the protein, Figure 3.5 shows a representative overlay of the ^{1}H-^{15}N HSQC spectrum of apo N-CaM at the low and higher ionic strength (I = 35 mM and 155 mM) at pH 6.4. Similar to the pH titration, the Δδ for residues in analogous positions were plotted as a function of ionic strength (Figure 3.6.A and 3.6.B). A general trend is displayed by the residues in both binding site I and II, such that no large fluctuations are observed.

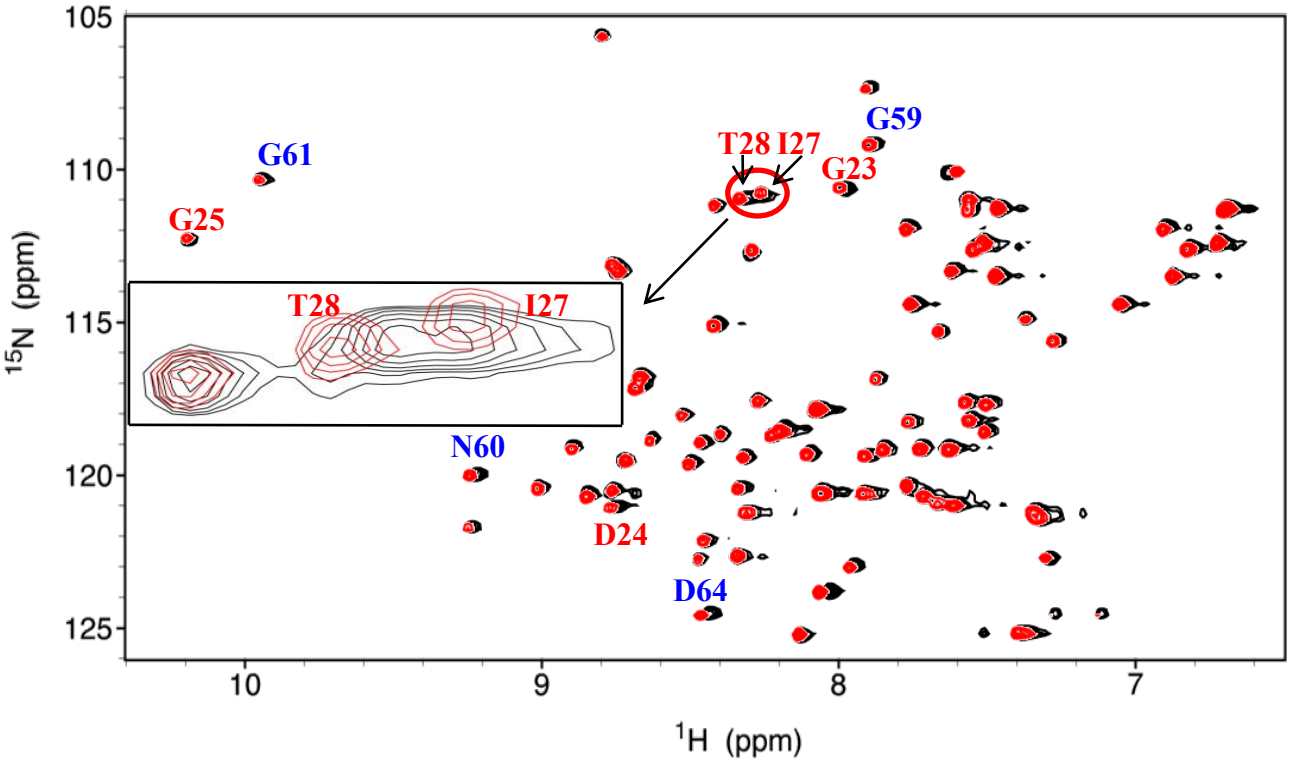

Figure 3.5 **Series of ^{1}H-^{15}N HSQC spectra from ionic strength titration at pH 6.4 of apo N-CaM**. Buffer conditions in 20 mM Mes, 10 mM EDTA at I = 35 mM (black) and I = 155 mM (red). Residues in Ca^{2+}-binding sites I (red) and II (blue). Inset displays peaks, T28 and I27, at the two ionic strengths. Data acquired on a Varian Inova 600 MHz at 25 °C.

The average of the maximum deviation from the original chemical shift position over the course of the titration for these residues in binding site I and II are 0.029 and 0.025, respectively. These deviations are small in comparison to the large shifts displayed in binding site II from the pH titration. Both sites appear largely unaffected by increasing ionic strength at pH 6.4.Based on

the appearance of peaks, the most obvious change appears to occur to the peak shapes of I27 and T28. These peaks are overlapped when the ionic strength is low (I = 35 and 65 mM), however between I = 95 and 125 mM the peaks become increasingly resolvable with increasing ionic strength (Figure 3.5, inset). At I = 155 mM, the peaks are distinctly separated and can be identified as belonging to residues I27 and T28 and the peak widths and positions easily determined.

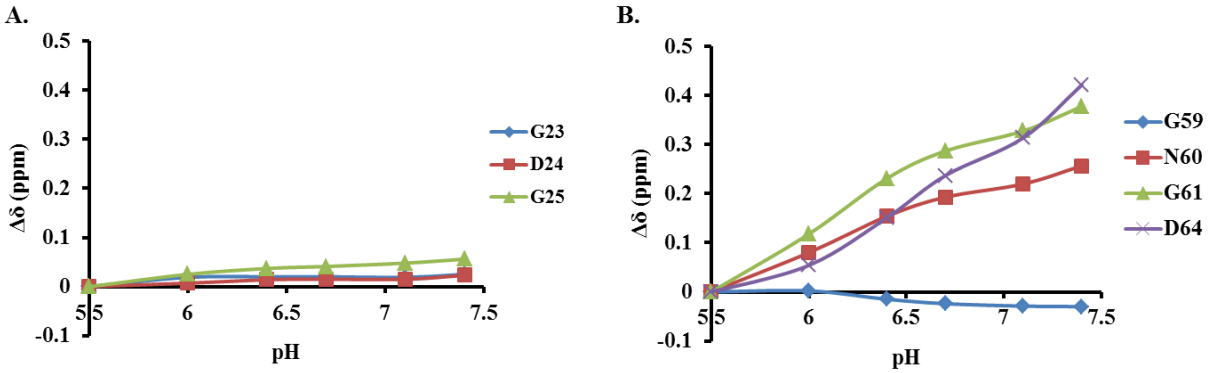

Figure 3.6 **Change in chemical shift position (Δδ) of residues from pH titration** The Δδ of residues in Ca^{2+}-binding loops, based on ^{1}H-^{15}N HSQC spectra from pH titration. **A.** Residues in Ca^{2+}-binding site I. **B.** Residues in Ca^{2+}-binding site II.

The ^{1}H - ^{15}N HSQC spectra from the ionic strength titration at pH 7.1 is displayed as an overlay of the spectra from I = 35 mM and I = 155 mM (Figure 3.7). Based on this overlay, more individual residues are perturbed over the duration of the ionic strength titration compared to the titration at pH 6.4. This spectrum shows broadening of the majority of the peaks and more peaks become shifted upfield. In binding site I, at I = 35 and 155 mM the peaks do not overlap completely for residues G23, D24, and G25, but instead small changes in peak positions occur. When comparing the residues of interest that were affected the most (GNG in site II) and the least (GDG in site I) by the pH titration, the same trend is observed from the ionic strength

titration at pH 7.1. The most pronounced changes in peak position occur to N60, G61, and D64 (Figure 3.7, insets). Figure 3.8 displays minimal changes to residues in site I ($\Delta\delta \leq 0.057$ ppm). In site II, a large shift in the $\Delta\delta$ occurs upon the addition of I = 65 mM for residues N60, G61, and D64 (insets). An increasing change in the chemical shift positions continue to occur to these three residues up to I = 95 mM, but no large fluctuations occur thereafter. Residue G59 is relatively unaffected by increasing I, based on the lack of observable change in chemical shift position. Similar to the ionic strength titration at pH 6.4, the peaks belonging to I27 and T28 become better resolved above I = 95 mM based on the sharper linewidths.

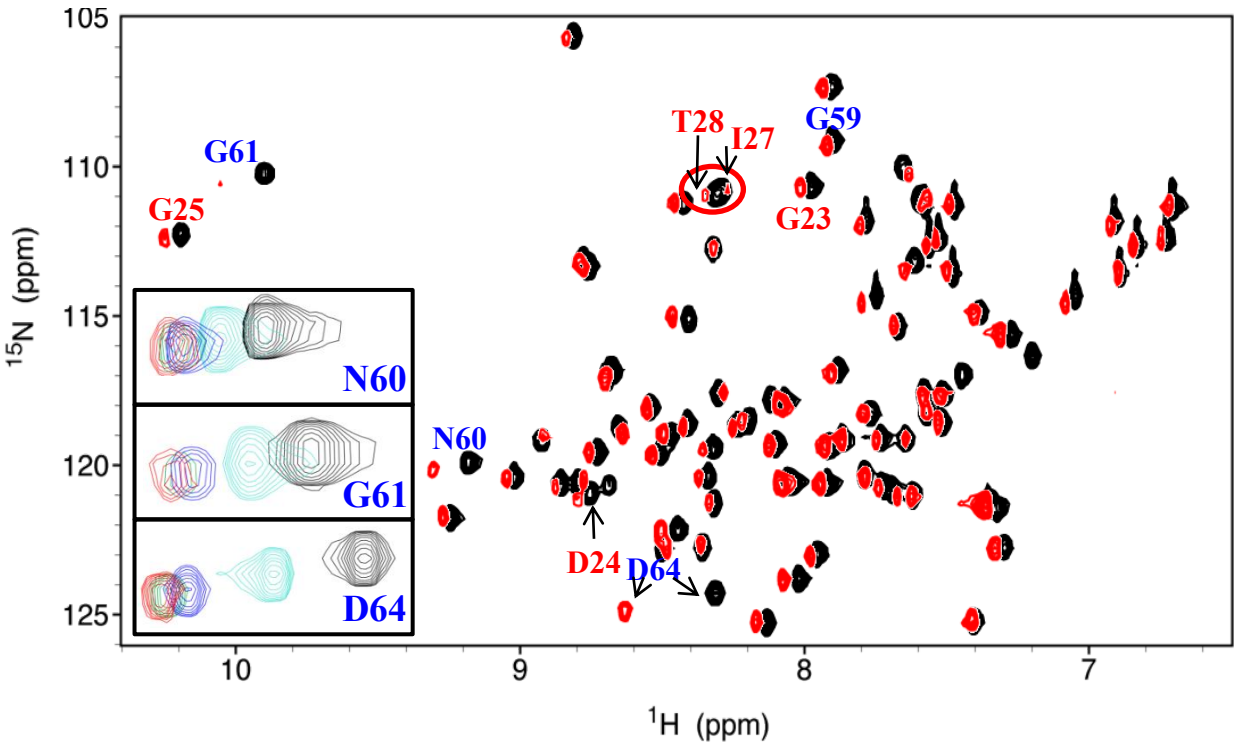

Figure 3.7 **A ^{1}H-^{15}N HSQC spectra of ionic strength titration at pH 7.1.** Apo N-CaM (1 mM) in 20 mM Hepes, 10 mM EDTA. I = 35 mM (black) and I = 155 mM (red). Residues in Ca^{2+} binding sites I (red) and II (blue). Insets: Residues undergoing change in peak position at each ionic strength: I = 65 mM (turquoise) I = 95 mM (blue) I = 125 mM (green) . Data acquired on a Bruker Avance 600 MHz at 25 °C.

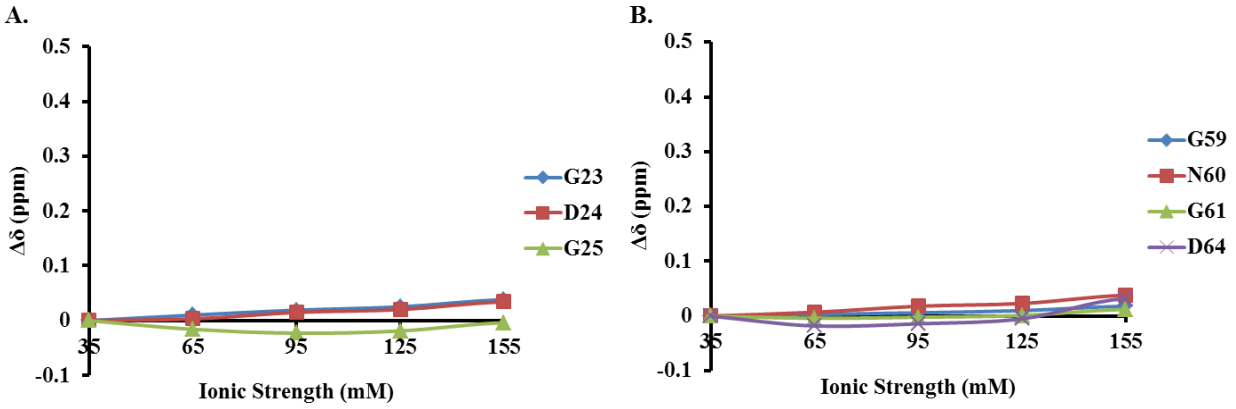

Figure 3.8 **Change in chemical shift position (Δδ) of residues from ionic strength titration at pH 6.4.** The Δδ of residues in Ca^{2+}-binding loops, based on ^{1}H-^{15}N HSQC spectra. **A.** Residues in Ca^{2+}-binding site I. **B.** Residues in Ca^{2+}-binding site II.

3.3.3 Thermal Stability of Apo N-CaM Evaluated by Far-UV CD

Far-UV circular dichroism (CD) spectroscopy was employed to evaluate the effect of pH and ionic strength on the thermal stability of the secondary structure of apo N-CaM. To observe alterations and loss of secondary structure induced by increasing temperature, the loss of the α-helical structure was monitored at 222 nm over the temperature range of 10 to 87.5 °C. The CD data from the temperature titration at pH 5.5, 6.4, and 7.1 were examined at equivalent ionic strength (I = 5 mM) (Figure 3.9) to evaluate the effect of varied pH on thermal stability. The CD signal at 222 nm at each pH displays a monophasic curve in which no transitions occur below 50 °C. Above 50 °C, a decrease in negative molar ellipticity begins to occur with increasing temperature for both pH 6.4 and 7.1. Based on the melting transition temperatures (T_m) acquired from the CD experiment (Table 3.1), apo N-CaM appears to become more susceptible to unfolding at lower temperatures with increasing pH compared to pH 5.5. The T_m of apo N-CaM is increased by approximately 10 °C between each pH used in this study. The lower T_m obtained (70.1 °C) at pH 6.4 and 7.1 (58.9 °C) confirm apo N-CaM is more thermally stable at pH 5.5 (80.6 °C).

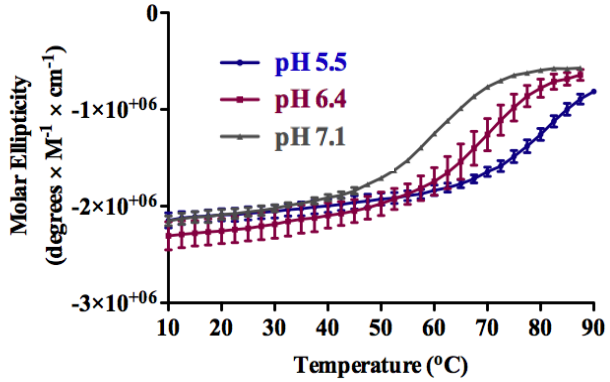

Table 3.1 Melting transition temperatures (T_m) obtained by CD spectroscopy in various pH conditions

pH	T_m (°C)	SE[a]
5.5	80.6	1.2
6.4	70.1	2.9
7.1	58.9	0.4

[a] standard error, where n = 2

Figure 3.9 **Thermal unfolding of apo N-CaM pH 5.5, 6.4, and 7.1**. CD signal monitored at 222 nm to assess physical alterations of the secondary structure in different pH conditions, pH 5.5, 6.4, and 7.1.

CD spectral melt data were then also acquired from the temperature titration over the ionic strength range of 5 – 125 mM at pH 6.4 and also at 7.1 (Figures 3.10.A and 3.10.B). Between *I* = 5 mM and 95 mM, the melting curves at each pH show a similar trend of increasing transition temperature with increasing ionic strength (Figure 3.11 and Table 3.2). At pH 6.4, the T_m remains relatively constant, staying within a narrow range between 70.1 °C and the maximum T_m, 76.6 °C, at *I* = 95 mM. The increase in ionic strength at *I* = 125 mM appears to have the opposite effect, and does not provide further thermal stabilization of apo N-CaM at pH 6.4 as determined by T_m.

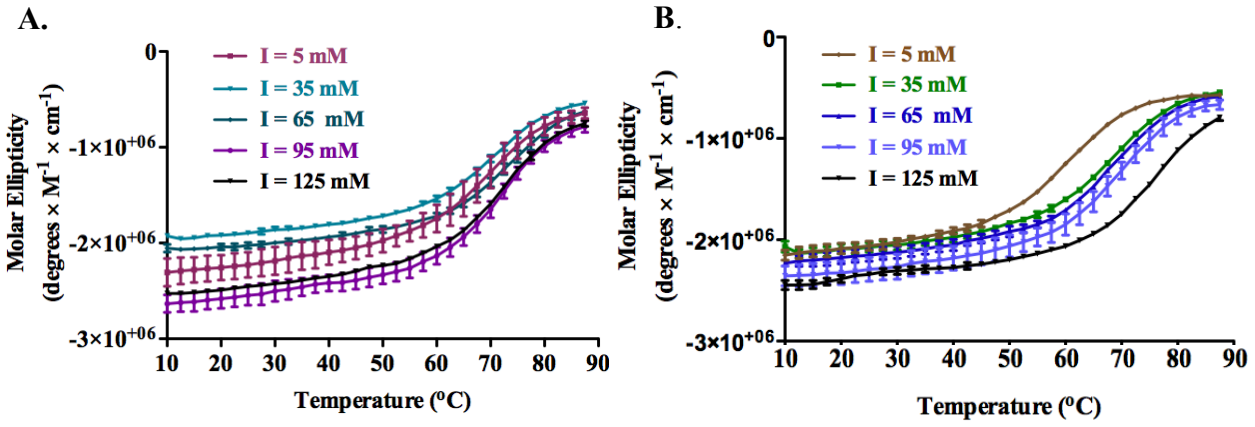

Figure 3.10 **Thermal unfolding of apo N-CaM in increasing ionic strength.** Physical alterations of the secondary structure were assessed in increasing ionic strength, *I* = 5 mM - 125 mM monitored by CD at 222 nm **A.** Ionic strength titration at pH 6.4. **B.** Ionic Strength titration at pH 7.4.

The curves, however, show that the overall amount of initial helical content is increased with increasing *I* (Figure 3.10). The same trends are observed with N-CaM at pH 7.1, but the degree to which the T_m and total α-helical content differ. A smaller change in total helicity is observed at lower temperatures but T_m is improved substantially by increasing *I*.

Based on the broader range of melting transition temperatures displayed in Table 3.2 and Figure 3.11, increasing ionic strength has a larger impact on the temperature titration at pH 7.1. The initial increase of ionic strength from *I* = 5 to 35 mM provides a 10-degree increase in the melting temperature. The impact to the T_m observed between 35 and 95 mM, however, is negligible. The largest incremental change in T_m is observed between *I* = 95 and 125 mM at pH 7.1, where the increase in T_m exceeds the highest T_m of both pH 6.4 (80.6 °C) and 5.5 (76.6 °C). Interestingly, apo N-CaM is much more sensitive to *I* at pH 7.1 than at 6.4 ($\Delta T_{m6.4}$ = 4 °C and $\Delta T_{m7.1}$ = 24 °C).

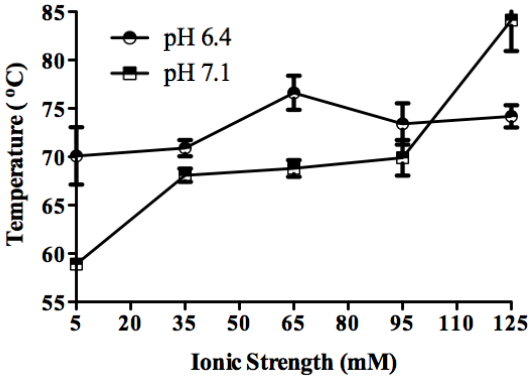

Table 3.2 Melting transition temperatures (T_m) obtained of apo N-CaM by CD spectroscopy in increasing ionic strength

Ionic Strength	pH 6.4		pH 7.1	
	Tm (°C)	SE[a]	Tm (°C)	SE[a]
5 mM	70.1	2.9	58.9	0.4
35 mM	70.9	0.8	68.1	0.7
65 mM	73.4	2.1	68.8	0.9
95 mM	76.6	1.8	69.9	1.8
125 mM	74.2	1.2	84.2	3.2

Figure 3.11 **The melting transition temperatures from the thermal unfolding of apo N-CaM at pH 6.4 and 7.1**. CD signal monitored at 222 nm to assess physical alterations of the secondary structure in increasing ionic strength.

3.4 Discussion

The apo form of the N-terminal domain of calmodulin was used as a model system to evaluate the usage of NMR as an investigative tool to assess the effect of solution conditions on protein stability. In this study, NMR spectroscopy was employed to detect structural perturbations occurring to individual residues as a result of pH and ionic strength titrations. A comparison of the spectra acquired from the 2D NMR experiments over the course of the pH titration displayed both large changes in δ for a subset of individual residues and small overall conformational adjustments. Individual residues that were greatly affected by the change in the solution environment were N60, G61, and D64 in site II. This was apparent from the obvious changes in the ^{1}H chemical shift positions of these residues relative to the modest overall structural change. Based on previous studies of chemical degradation in CaM, it is known that N60 and G61 are positioned in a labile sequence (GNG) prone to deamidation (*8*). In the case of N-CaM, deamidation of the N60 residue in binding site II occurs relatively rapidly at neutral to basic pH (*9*). In the analogous positions, at the end of the binding loop of site I, a GDG sequence exists where residue D24 is known to undergo isomerization at neutral pH, albeit at a slower rate than deamidation of N60 (*8*). In our studies the data was collected in a manner to minimize chmical degradation so that only physical changes in the protein would be observed. No large changes in the ^{1}H chemical shift positions occur to the residues in the GDG sequence, indicating the change in the solution environment does not affect these residues while GNG in site II is substantially altered, despite both regions having the same structure.

The Ca^{2+} binding sites I and II are structurally equivalent, based on the NMR structures determined by Ikura *et al*. and crystallographic structure by Babu *et al*. (*43*, *44*). The secondary structural elements contained within the Ca^{2+} binding sites is a loop, a short antiparallel β-strand

proceeded by a helix. The two sites adjoin one another and are connected by hydrogen-bonding between their β-strand. The primary sequences are comprised of the same residues at each position that coordinate to the Ca^{2+} ion, with the exception of position 3 in the loop and position 5 in the β-turn. At position 3, in the absence of calcium in many studies has confirmed that CaM is highly susceptible to deamidation and/or isomerization at neutral to basic pH (*9*, *23*, *24*). This study is the first to use NMR spectroscopy to identify the specific residues affected by changing solution conditions to determine how protein stability is altered by its environment. It is interesting that residues in structurally equivalent locations are affected differently, particularly because the sequence identity is high (58%) and these residues are surface exposed. As such, it is expected that D24 would be in the anionic state throughout the pH titration. N-CaM has no histidine (pKa ~ 6), making it unclear why the structure of site I would differ from site II.

A broader pH titration spanning over pH 3 to pH 12 could affect free ionizable side-chains, such as aspartate (Asp) and glutamate (Glu) located in the binding loops. The pH titration in this study occurs outside the range of side-chain pKa's. Residue D64 most likely remains ionized during this titration and is not a site where isomerization is known to occur. It is unknown what would cause the dramatic shifting in chemical shift position of D64 without additional experiments.

Further analysis of the peak shape changes occurring over the course of the pH titration can provide information about whether higher pH is affecting the flexibility of N-CaM. Evaluating the sharpening or broadening of the peak shape provides a rapid way to estimate dynamic changes in the protein in the different solution conditions. Based on previous NMR relaxation measurements and structural studies, the backbone of apo CaM is highly flexible in the binding loop region of the Ca^{2+} binding sites I (residues 21-30) and II (residues 56-67) (*45*).

The spectra acquired from the pH titration, do not display a uniform sharpening of the peaks with increasing pH. Instead, the specific residues undergoing dramatic chemical shift changes (N60, G61, and D64) become sharper while other unaltered residues remain relatively broad. The sharpening of the individual peaks suggests this portion of the protein becomes more flexible with increasing pH. These individual residues are experiencing more movement in solution and the peaks begin to sharpen because the exchange between the conformations they are undergoing is relatively fast compared to the NMR timescale. The majority of the peaks remain broad due to conformational averaging from the relatively slower exchange of conformations the protein is undergoing compared to the NMR timescale (*28*). Backbone flexibility in the binding loop region and high pH could be the major contributors to pronounce changes observed in the ^{1}H chemical shift positions of N60 and G61 at higher pH, and this local change in conformational flexibility correlates with decreased thermal stability for N-CaM. Based on the spectra acquired from the pH titration, global structure is retained throughout the titration, but local structure at binding site II is altered to a higher degree, while binding site I is relatively unaffected.

To better understand if the impact of local structure at each pH is a consequence of overall physical stability, a temperature titration was performed and monitored by far-UV CD. Previous studies have reported similar melting transition temperatures of isolated apo N-CaM at pH 6.8 and 8.0, resulting from temperature titrations monitored by CD, fluorescence, and 2D NMR experiments (*9, 19*). Apo N-CaM unfolds at higher temperatures compared to these studies of the isolated domain, however, this may be due to the absence of a chemical denaturant. Based on the molar ellipticity at 10 °C, the same amount of helicity is retained upon incubation at each pH before the temperature titration, indicating pH is not affecting secondary structure initially. Over the course of the temperature titration, the unfolding of apo N-CaM, however,

occurs at lower temperatures with increasing pH. The loss of the secondary structure as a function of temperature appears to be impacted at each pH. Although CD is unable to detect local perturbations, the NMR spectra reveal perturbations are occurring at individual residues, particularly in site II, upon titration from pH 5.5 to 7.4. Based on the CD data, the physical stability of N-CaM becomes more thermally labile with increasing pH and when coupled together with the NMR data, they indicate that physical instability of apo N-CaM at higher pH is likely due to alterations in structure at site II.

Along with identifying residues that are individually affected by destabilizing solution conditions, 2D NMR experiments were used to identify residues impacted by known stabilizing conditions. Masino *et al*., reported high ionic strength (≥100 mM KCl) plays a role in thermally stabilizing the secondary and tertiary structure of CaM at pH 8. Pharmaceutical formulations of proteins include ionic compounds in the buffer components to stabilize proteins in aqueous solutions (*16*). Many ions are used in these formulations to increase solubility and thermal stability of a protein in aqueous solutions (*3, 16*). Various reviews and studies propose explanations of how ionic compounds, specifically anions, can affect the stability of proteins (*18, 46*). One common explanation is based on the trend observed of how proteins are affected by anions in the Hofmeister series (*47*). The Hofmeister series ranks anions according to on their ability to stabilize (salt-in) or destabilize (salt-out) proteins in moderate to high salt concentrations (0.01 – 1 M) (*48*). The mechanism of stabilization or destabilization of a protein upon the addition of the Hofmeister series is still not well-understood. Included in the Hofmeister series is the Cl^- ion, which has a neutral effect with respect to stability. Because the Cl^- ion has a neutral effect on stability, the Hofmeister effect cannot be used to explain how KCl thermally stabilizes N-CaM.

The ionic strength titration using KCl was performed at pH 6.4 and 7.1 because rapid spontaneous degradation via deamidation/isomerization to proteins is known to occur above pH 6 (*2, 16*). Based on the spectra, the overall structure and specific residues at pH 6.4 and 7.1 were impacted differently by the increasing ionic strength. No large movements in chemical shift positions of individual residues were observed indicating the conformation of N-CaM at pH 6.4 remained unaltered by the presence of high ionic strength. At pH 7.1, however, the conformation is altered slightly based on the shift of the majority of peaks. Additionally, larger chemical shift changes at specific residues within binding site I indicate the structure is not uniformly affected by ionic strength at pH 7.1.

One explanation, which can affect specific sites within N-CaM, is the alteration to the solvent dielectric upon increasing the ionic strength. Increasing the solvent dielectric constant, could impact electrostatic and hydrophobic interactions between residues. Increasing dielectric constant can drive hydrophobic residues to become buried within the core and less solvent exposed. Disruption of electrostatic interactions within macromolecular complexes could also be a consequence of adding ions to the solution. Changing the physical environment by increasing the solvent dielectric is known to affect deamidation rates. A higher dielectric constant has the ability to stabilize the ionic intermediate, thereby accelerating the rate of deamidation of labile aparagines (*16*). Based on the spectra, residues exposed in the binding site and located in the deamidation sequence, are moving downfield indicating the ionic strength is directly affecting local structure to a higher degree than the whole structure. Although more extensive studies required to ascertain why certain residues are impacted while others are not, it is interesting that some of the residues that are adjacent or reside in labile deamidation and isomerization sequences, are the same residues impacted by the pH and ionic strength titrations.

Although specific sites are impacted substantially at pH 7.1 compared to pH 6.4, CD data suggests the overall structure of apo N-CaM at each pH is thermally stabilized by high ionic strength. Secondary structure seems initially to be more impacted at pH 6.4 compared to pH 7.1, however each pH melt follow the same trend in which each curve displays an initial increase in helicity (beginning at 10 °C) with increasing ionic strength. There appears to be a relationship between increased helicity in high ionic strength and the ability of ionic strength to thermally stabilize apo N-CaM at pH 6.4 and pH 7.1. Conclusively, the increase in ionic strength appears to stabilize the secondary structure, as indicated by increased T_m within this pH range.

Analysis of the NMR peak shapes at pH 6.4 and 7.1, reveals uniform peak sharpening occurs as ionic strength is increased. This peak sharpening indicates apo N-CaM becomes more rigid with increasing ionic strength. The peak positions of many residues have also changed upon increasing ionic strength at pH 7.1 compared to the ionic strength titration at pH 6.4. Also, the same pH-sensitive residues in Ca^{2+}-binding site II (N60, G61, D64) undergo more movement compared to the residues in binding site I. The addition of increased ionic strength to the higher pH environment creates an additive effect on the mobility and physical perturbation of these residues in binding site II. At specific sites (I27/T28) the change in peak shape suggests the dynamics of local structure is affected increasingly with ionic strength while the overall structure becomes less dynamic. Altogether with the pH titration, our data suggests thermal stabilization from high ionic strength is related to a combination of local structure interactions and overall protein dynamics.

3.5 Summary

In this study, NMR spectroscopy was used to obtain site-specific information to assess the structural stability of the model system, apo N-CaM, in different solution conditions known to affect stability. The ^{1}H-^{15}N NMR experiments were used as an investigative tool to acquire molecular information about the conformational changes and residues perturbed upon increasing pH and ionic strength. The data from a series of spectra collected during pH and ionic strength titrations show residues in highly similar sequences and within analogous structural elements are impacted differently by their solution environment. As such, it is evident that conditions affect local structure. In N-CaM these perturbations occur to residues located near a sequence prone to deamidation. Because the NMR data was collected before deamidation products were observed in the spectra, the data reflect only physical changes to N-CaM

The data derived from the NMR experiments and the thermal titration were used to evaluate the overall physical stability of apo N-CaM in different solution conditions in order to explore the relationship between local structure and overall thermal stability. A combination of low- and high-resolution techniques contributed to establishing that specific residues are affected by alterations in the solution environment. Based on alterations to the thermal stability and perturbation of specific residues in these conditions, it appears that overall thermal stability of a protein depends on a subset of individual residues that are modulated by solution conditions.

3.6 References

1. Bartlett, A. I., and Radford, S. E. (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms, *Nat Struct Mol Biol 16*, 582-588.
2. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W., and Katayama, D. S. (2010) Stability of protein pharmaceuticals: an update, *Pharm Res 27*, 544-575.
3. Wang, W. (1999) Instability, stabilization, and formulation of liquid protein pharmaceuticals, *Int J Pharm 185*, 129-188.
4. Reubsaet, J. L., Beijnen, J. H., Bult, A., van Maanen, R. J., Marchal, J. A., and Underberg, W. J. (1998) Analytical techniques used to study the degradation of proteins and peptides: physical instability, *J Pharm Biomed Anal 17*, 979-984.
5. Reubsaet, J. L., Beijnen, J. H., Bult, A., van Maanen, R. J., Marchal, J. A., and Underberg, W. J. (1998) Analytical techniques used to study the degradation of proteins and peptides: chemical instability, *J Pharm Biomed Anal 17*, 955-978.
6. Skinner, A. L., and Laurence, J. S. (2010) Probing Residue-Specific Interactions in the Stabilization of Proteins Using High-Resolution NMR: A Study of Disulfide Bond Compensation, *Journal of Pharmaceutical Sciences 99*, 2643-2654.
7. Siuzdak, G. (2006) *The Expanding Role of Mass Spectrometry in Biotechnology* 2ed., MCC Press, San Diego.
8. Aswad, D. W., (Ed.) (1995) *Deamidation and isoaspartate formation in peptides and proteins*, CRC Press.
9. Potter, S. M., Henzel, W. J., and Aswad, D. W. (1993) In vitro aging of calmodulin generates isoaspartate at multiple Asn-Gly and Asp-Gly sites in calcium-binding domains II, III, and IV, *Protein Sci 2*, 1648-1663.
10. Robinson, N. E., and Robinson, A. B. (2004) Molecular Clocks: Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins., Althouse Press, Cave Junction.
11. Li, Y., Williams, T. D., Schowen, R. L., and Topp, E. M. (2007) Trehalose and Calcium Exert Site-Specific Effects on Calmoduin Conformation in Amorphous Solids *Biotechnology and Bioengineering 97*, 4.
12. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism, *Bba-Proteins Proteom 1751*, 119-139.
13. Esfandiary, R., Hunjan, J. S., Lushington, G. H., Joshi, S. B., and Middaugh, C. R. (2009) Temperature dependent 2nd derivative absorbance spectroscopy of aromatic amino acids as a probe of protein dynamics, *Protein Sci 18*, 2603-2614.
14. Wang, W., and Roberts, C. J., (Eds.) (2010) *Aggregation of Therapeutic Proteins*, WILEY.
15. Pearlman, R., and Bewley, T. A., (Eds.) (1993) *Stability and Characterization of Human Growth Hormone*, Plenum Press, New York.
16. Wakankar, A. A., and Borchardt, R. T. (2006) Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization, *J Pharm Sci 95*, 2321-2336.

17. Stratton, L. P., Kelly, R. M., Rowe, J., Shively, J. E., Smith, D. D., Carpenter, J. F., and Manning, M. C. (2001) Controlling deamidation rates in a model peptide: effects of temperature, peptide concentration, and additives, *J Pharm Sci 90*, 2141-2148.
18. Baldwin, R. L. (1996) How Hofmeister Ion Interactions affect Protein Stability *Biophysical Journal 71*.
19. Biekofsky, R. R., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1998) Ca2+ coordination to backbone carbonyl oxygen atoms in calmodulin and other EF-hand proteins: 15N chemical shifts as probes for monitoring individual-site Ca2+ coordination, *Biochemistry 37*, 7617-7629.
20. Jurado, L. A., Chockalingam, P. S., and Jarrett, H. W. (1999) Apocalmodulin, *Physiol Rev 79*, 661-682.
21. Klee, C. B., Crouch, T. H., and Richman, P. G. (1980) Calmodulin, *Annu Rev Biochem 49*, 489-515.
22. Mal, T. K., and Ikura, M. (2006) NMR Investigation of Calmodulin, *Modern Magnetic Resonance*.
23. Ota, I. M., and Clarke, S. (1989) Calcium affects the spontaneous degradation of aspartyl/asparaginyl residues in calmodulin, *Biochemistry 28*, 4020-4027.
24. Masino, L., Martin, S. R., and Bayley, P. M. (2000) Ligand binding and thermodynamic stability of a multidomain protein, calmodulin, *Protein Sci 9*, 1519-1529.
25. Kasimova, M. R., Kristensen, S. M., Howe, P. W., Christensen, T., Matthiesen, F., Petersen, J., Sorensen, H. H., and Led, J. J. (2002) NMR studies of the backbone flexibility and structure of human growth hormone: a comparison of high and low pH conformations, *J Mol Biol 318*, 679-695.
26. Ricci, M. S., and Brems, D. N. (2004) Common structural stability properties of 4-helical bundle cytokines: possible physiological and pharmaceutical consequences, *Curr Pharm Des 10*, 3901-3911.
27. Wells, J. A., and de Vos, A. M. (1993) Structure and Function of Human Growth Hormone: Implications for the Hematopoietins, *Annual Review of Biophysics and Biomolecular Structure 22*, 329-351.
28. Skinner, A. L., and Laurence, J. S. (2008) High-field solution NMR spectroscopy as a tool for assessing protein interactions with small molecule ligands, *J Pharm Sci 97*, 4670-4695.
29. Finn, B. E., and Forsen, S. (1995) The evolving model of calmodulin structure, function and activation, *Structure 3*, 7-11.
30. Ikura, M. (1996) Calcium binding and conformational response in EF-hand proteins, *Trends Biochem Sci 21*, 14-17.
31. Yang, J. J., Gawthrop, A., and Ye, Y. (2003) Obtaining site-specific calcium-binding affinities of calmodulin, *Protein Pept Lett 10*, 331-345.
32. Sorensen, B. R., and Shea, M. A. (1998) Interactions between domains of apo calmodulin alter calcium binding and stability, *Biochemistry 37*, 4244-4253.
33. Weickert, M. J., Doherty, D. H., Best, E. A., and Olins, P. O. (1996) Optimization of heterologous protein production in Escherichia coli, *Curr Opin Biotechnol 7*, 494-499.
34. Bartlett, R. K., Urbauer, R. J., Anbanandam, A., Smallwood, H. S., Urbauer, J. L., and Squier, T. C. (2003) Oxidation of Met144 and Met145 in calmodulin blocks calmodulin dependent activation of the plasma membrane Ca-ATPase., *Biochemistry* 9.
35. Schägger, H. (2006) Tricine-SDS-PAGE, *Nature Protocols*, 6.

36. Masino, L., Martin, S. R., and Bayley, P. M. (2000) Ligand Binding and thermodynamic stability of a multidomain protein, calmodulin, *Protein Science 9* 10.
37. Carafoli, E. (1987) Intracellular calcium homeostasis, *Annu Rev Biochem 56*, 395-433.
38. Balog, E. M., Lockamy, E. L., Thomas, D. D., and Ferrington, D. A. (2009) Site-specific methionine oxidation initiates calmodulin degradation by the 20S proteasome, *Biochemistry 48*, 3005-3016.
39. Tarcsa, E., Szymanska, G., Lecker, S., O'Connor, C. M., and Goldberg, A. L. (2000) Ca2+-free calmodulin and calmodulin damaged by in vitro aging are selectively degraded by 26 S proteasomes without ubiquitination, *J Biol Chem 275*, 20295-20301.
40. Chou, J. J., Li, S., Klee, C. B., and Bax, A. (2001) Solution structure of Ca(2+)-calmodulin reveals flexible hand-like properties of its domains, *Nat Struct Biol 8*, 990-997.
41. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR, *J Biomol NMR 6*, 135-140.
42. Goddard, T. D., and Kneller, D. G. (2004) SPARKY, San Francisco: University of California.
43. Ikura, M., Kay, L. E., and Bax, A. (1990) A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin, *Biochemistry 29*, 4659-4667.
44. Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R., and Cook, W. J. (1985) Three-dimensional structure of calmodulin, *Nature 315*, 37-40.
45. Finn, B. E., Evenas, J., Drakenberg, T., Waltho, J. P., Thulin, E., and Forsen, S. (1995) Calcium-induced structural changes and domain autonomy in calmodulin, *Nat Struct Biol 2*, 777-783.
46. Sedlak, E., Stagg, L., and Wittung-Stafshede, P. (2008) Effect of Hofmeister ions on protein thermal stability: roles of ion hydration and peptide groups?, *Arch Biochem Biophys 479*, 69-73.
47. Zhang, Y., and Cremer, P. S. (2006) Interactions between macromolecules and ions: the hofmeister series, *Current Opinion in Chemical Biology 10*, 658-663.
48. Creighton, T. E., (Ed.) (1993) *Proteins: Structures and Molecular Properties*, 2nd ed., W.H. Freeman and Company, New York.

Page Left Intentionally Blank

CHAPTER 4

SUMMARY

4.1 Summary

The work described in this dissertation demonstrates how to maximize production of the model system, N-CaM using the ligand protectant, calcium, in the expression media to enhance *in vivo* stability. The high-yield of N-CaM, acquired from the expression and purification, facilitated the overall goal of developing a method using NMR spectroscopy to evaluate protein stability in different solution conditions. Site-specific information about the stability of N-CaM was obtained by performing 2D NMR experiments to probe local interactions and global structure in solution conditions known to affect thermal stability. This data provides important information that can facilitate our understanding of the relationship between overall structural stability and specific residues undergoing physical alterations as a consequence of local environment perturbations. Our approach to obtaining molecular information to understand local and global stability of a protein can contribute to research studying the structural stability of proteins with therapeutic interest.

4.1.1 Using NMR spectroscopy to probe structural stability

NMR spectroscopy was employed to obtain molecular information to assess global conformational changes and structural perturbations of specific residues in the model system N-CaM in increasing pH and ionic strength. The current methodologies used to study the structural stability of proteins involve investigating the global structure by spectroscopy and calorimetry to acquire thermal stability. Site-specific information is not attained from these techniques and, therefore, any local perturbation that may impact global stability is not considered. Our studies

demonstrated how NMR spectroscopy can be used to reveal differences in local and global structure upon changing the solution environment. Specifically, NMR spectroscopy provided a way to investigate how residues in structurally equivalent sites are affected differently in increasing pH and ionic strength. Identifying the residues largely affected in these conditions has shown local instability exists when minimal structural stability is observed. This observation is only revealed when a technique, such as NMR, is able to detect subtle perturbations and simultaneously global structural changes. Molecular information acquired in this methodology can be used to facilitate the understanding of the mechanisms of physical and chemical stability and the environmental factors that influence the instability of a protein.

We have provided a methodology employing NMR spectroscopy to obtain molecular information that can also facilitate formulating therapeutic proteins in stabilizing conditions. Site-directed mutagenesis, varying solution conditions, and excipient screenings are common approaches to finding the optimal stable formulation. Utilizing NMR to identify residues involved in local instability can be used in combination other spectroscopic and calorimetric tools to identify conditions and/or mutations that contribute to stabilizing the overall structural stability of proteins. Alternatively, NMR spectroscopy can be employed in studies interested in understanding the mechanism of degradation via "hot spots" in proteins. Regions where "hot spots" exists contain residues in labile sequences which are prone to chemical degradation, such as deamidation and/or isomerization, and can have structural consequences as well. Numerous applications of NMR experiments can facilitate research interested in obtaining residue-specific information to understand mechanisms of degradation and stabilization, and is therefore, invaluable to studying and designing safe and effective protein based therapeutics.